

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE INFORMÁTICA



## **Aprendizagem automática para classificar a cobertura do solo em espaços verdes urbanos através de dados de satélite**

Adriano António Pinheiro Mão de Ferro

**Mestrado em Ciência de Dados**

Dissertação orientada por:  
Professor Doutor Pedro Ferreira  
Professora Doutora Sara Madeira



*“Machine learning will bring about not just a new era of civilization, but a new stage in the evolution of life on earth.”*

— Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*



# Agradecimentos

Ao longo da minha vida, que já não é curta, fui consolidando a ideia de que a educação e a aprendizagem são fundamentais para a completitude do ser humano e é por isso que, de forma um pouco improvável ou pouco expetável, estou a apresentar esta dissertação de mestrado numa área um tanto à margem daquela em que sempre desenvolvi a minha atividade profissional e da minha formação inicial, em Economia. Tal não seria possível sem a qualidade do sistema de ensino português e é por isso que começo por agradecer à Universidade de Lisboa e em particular à Faculdade de Ciências esta oportunidade que me foi dada. Mas como as pessoas é que fazem as instituições, estendo estes agradecimentos aos meus orientadores, Prof<sup>a</sup> Sara Madeira e Prof. Pedro Ferreira pelo suporte pedagógico, científico e técnico e pela disponibilidade que me dispensaram no particular contexto da pandemia de COVID-19 em que este trabalho decorreu.

Tendo em conta a área de estudo desta dissertação, os dados são o principal ingrediente do trabalho desenvolvido, por isso quero agradecer à doutoranda Cristiana Aleixo e ao Prof. Pedro Pinho do Departamento de Biologia Vegetal da FCUL pela disponibilização dos dados e pelos esclarecimentos fundamentais para o seu tratamento e ao entendimento do problema em estudo.

À minha mulher, Lina, agradeço o encorajamento, o suporte emocional e a inabalável confiança nas minhas capacidades com que sempre me distinguiu.

Aos meus filhos Ana e Carlos, para além do óbvio apoio afetivo, agradeço o seu contributo objetivo para o sucesso deste trabalho, partilhando a sua experiência, os seus conhecimentos específicos e as suas sugestões estéticas e metodológicas.

Não posso também deixar de agradecer aos meus colegas de curso, em particular à Maria Fidalgo, João Lobo, Miguel Silva e Rodrigo Neves, pelo são convívio que me proporcionaram e também pelo apoio técnico que em alguns momentos foi para mim fundamental.



# Resumo

Atualmente, a maioria da população mundial vive em cidades. A qualidade de vida nas cidades é fortemente afetada pelo planeamento urbano e particularmente, pela existência de Espaços Verdes Urbanos (EVU). A inventariação destes espaços é, por isso, da maior importância. As imagens da superfície terrestre obtidas por deteção remota, através de satélites artificiais, permitem identificar a cobertura do solo através do espectro eletromagnético. Esta é, no entanto uma tarefa complexa, devido à diversidade das características envolvidas (árvores e arbustos, suas dimensões, o tipo de folhagem, lagos e respetivas margens) e à variabilidade temporal das imagens em função das estações do ano. Algoritmos de aprendizagem automática para classificação de objetos através de imagens de satélite, como por exemplo, *Support Vector Machines* (SVM), árvores de decisão, redes neuronais, têm sido amplamente utilizados na análise da cobertura e uso do solo o que sugere a pertinência da sua utilização para classificar a composição dos EVU, em particular, Redes Neuronais Convolucionais (RNC) aplicadas a imagens de satélite.

O objetivo desta dissertação é verificar a eficácia de métodos de aprendizagem automática para classificar a cobertura do solo em EVU, a partir de dados extraídos de imagens de satélite. Foi utilizado um conjunto de dados pré-classificados obtidos no âmbito do projeto Bioveins [3]. Este projeto é uma parceria entre várias Universidades Europeias e a *BiodivERSA - the network programming and funding research on biodiversity and ecosystem services across European countries and territories*. A Universidade de Lisboa é uma destas Universidades e por isso os dados foram extraídos e preparados pelo Departamento de Biologia Vegetal da FCUL. O conjunto de dados engloba 220 EVU em 7 cidades europeias (Lisboa, Almada, Zurique, Paris, Antuérpia, Poznan e Tartu). Os dados iniciais estão estruturados em três grupos. O primeiro grupo inclui as coordenadas geográficas do centróide do espaço de 10 metros por 10 metros que representa cada instância. O segundo grupo inclui 27 características obtidas das imagens de satélite, nomeadamente as frequências do espectro eletromagnético visível e invisível ao olho humano, e um conjunto de índices de vegetação, solo, água, densidade urbana e outras variáveis biofísicas. Finalmente o terceiro grupo, corresponde à classificação dos componentes dos espaços, efetuada por observação no terreno. Esta classificação foi efetuada atribuindo uma percentagem a cada uma de 11 categorias: *Allotment gardens*, *Artificial surfaces*, *Ornamental bed*, *Bare-lands*, *Grassland*, *Water surfaces*, *Riparian vegetation*, *Shrubs*, *Broad deciduous*, *Broad ever* e *Coniferous*.

Foram efetuadas três abordagens de aprendizagem automática. Numa primeira abordagem ao problema, aplicámos um método *ensemble*, floresta aleatória, criando um classificador binário para cada categoria. Para esta abordagem foram utilizados como unidade de entrada, os centróides dos espaços,

sem qualquer transformação estrutural, mas o conjunto de dados foi transformado em 11 réplicas, uma para cada categoria, em que a cada instância, foi atribuída uma classe com o valor 0 ou 1, correspondendo à ausência ou presença da respectiva categoria. Na segunda abordagem, utilizámos um único classificador multi-classe, constituído por uma RNC. Os dados foram agrupados de forma a constituírem imagens para input da rede. A cada imagem foi atribuída uma classe através de um número inteiro correspondendo a uma das categorias, aquela com maior predominância nessa imagem, considerando todos os pontos que a formaram. A última abordagem baseou-se também numa RNC mas utilizando tantos classificadores quantas as categorias existentes. À semelhança do que foi feito para o método *ensemble*, cada imagem foi classificada com valores de 0 ou 1, correspondendo à ausência ou presença da categoria respectiva na imagem.

A classificação binária pelo método *ensemble* não produziu resultados muito satisfatórios, mas ambas as abordagens com RNC evidenciaram serem bastante eficazes para a classificação dos EVU. A abordagem multi-classe apresenta vantagens pela simplicidade de implementação pois utiliza um único classificador.

Os dados utilizados para aprendizagem apresentam um elevado desequilíbrio de classes, por isso a obtenção de mais dados e mais variados, poderá representar um desafio para futuro trabalho sobre o tema, assim como uma abordagem *multi-label* que pode permitir um resultado mais detalhado para cada espaço.

**Palavras Chave:** Espaços Verdes Urbanos, Classificação de Imagens de Satélite, Aprendizagem Automática, Redes Neurais Convolucionais, Aprendizagem Profunda.



# Abstract

Currently, the majority of the world's population lives in cities. The quality of life in cities is strongly affected by urban planning and particularly, by the existence of urban green spaces (UGS). The inventory of these spaces is, therefore, of the utmost importance. The images of the terrestrial surface obtained by remote sensing, through artificial satellites, allow to identify the ground cover through the electromagnetic spectrum. This is, however, a complex task, due to the diversity of the characteristics involved (trees and shrubs, their dimensions, the type of foliage, lakes and respective banks) and the temporal variability of images, according to the seasons. Machine Learning (ML) algorithms for object classification using satellite images, such as Support Vector Machine (SVM), Decision Tree (DT), Neural Network (NN), have been widely used in the analysis of coverage and land use, which suggests the relevance of its use to classify the composition of UGS, in particular, Convolutional Neural Networks (CNN) applied to satellite images. The objective of this dissertation is to verify the effectiveness of ML methods to classify the soil cover in UGS, based on data extracted from satellite images. A set of pre-classified data obtained within the scope of the project Bioveins [3] was used. This project is a partnership between several European Universities and *BiodivERsA the network programming and funding research on biodiversity and ecosystem services across European countries and territories*. The University of Lisbon is one of these Universities and therefore the data were extracted and prepared by the Department of Plant Biology at FCUL. The data set comprises 220 UGS in 7 European cities (Lisbon, Almada, Zurich, Paris, Antwerp, Poznan and Tartu). The initial data are structured in three groups. The first group includes the geographical coordinates of the centroid of the 10 meter by 10 meter space that represents each instance. The second group includes 27 characteristics obtained from satellite images, namely the frequencies of the electromagnetic spectrum visible and invisible to the human eye and a set of indices of vegetation, soil, water, urban density and other biophysical variables. Finally, the third group corresponds to the classification of the components of the spaces, carried out by observation on the ground. This classification was made by assigning a percentage to each of 11 categories, *Allotment gardens, Artificial surfaces, Ornamental bed, Bare-lands, Grassland, Water surfaces, Riparian vegetation, Shrubs, Broad deciduous, Broad ever and Coniferous*. Three machine learning approaches were followed. In a first approach to the problem, we applied an *ensemble* method, random forest, creating a binary classifier for each category. For this approach, the centroid of the spaces were used as an input unit, without any structural transformation, but the data set was transformed into 11 replicates, one for each category, in which each instance was assigned a class with the value 0 or 1, corresponding to the absence or presence of the respective

category. In the second approach, we used a single multiclass classifier, consisting of a CNN. The data were grouped in order to constitute images for network input. Each image was assigned a class through an integer corresponding to one of the categories, the one with the most predominance in that image, considering all the points that formed it. The latter approach was also based on a Convolutional Neural Network but using as many classifiers as the number of existing categories. Similar to what was done for the ensemble method, each image was classified with values of 0 or 1, corresponding to the absence or presence of the category respective image. The binary classification by the ensemble method did not produce very satisfactory results, but both approaches with CNN proved to be quite effective for the classification of UGS. The multiclass approach has advantages due to its simplicity of implementation, since it uses a single classifier. The data used for learning has a high class imbalance, so obtaining more and more varied data may represent a challenge for future work on the topic, as well as a multi-label approach that can allow a more detailed result for each space.

**Keywords:** Urban Green Spaces, Satellite Image Classification, Machine Learning, Convolutional Neural Network, Deep Learning.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Enquadramento e motivação	1
1.2	Objectivos e resultados esperados	2
1.3	Contribuições	2
1.4	Estrutura do documento	3
<b>2</b>	<b>Enquadramento e trabalho relacionado</b>	<b>5</b>
2.1	Descrição do problema	5
2.2	Algoritmos de aprendizagem automática para classificação	6
2.2.1	SVM	7
2.2.2	Árvores de decisão	7
2.2.3	Redes neuronais artificiais	8
2.2.4	Redes neuronais profundas	8
2.2.5	Métodos <i>ensemble</i>	9
2.2.6	Floresta aleatória	10
2.2.7	Extreme gradient boosting	10
2.2.8	Componentes de aprendizagem profunda: <i>Stride</i> e <i>Padding</i>	10
2.2.9	Escolha de parâmetros: <i>GridSearchCV</i>	12
2.3	Avaliação e seleção dos modelos	12
2.3.1	Métricas para avaliar o desempenho dos classificadores	13
2.3.2	O problema do desequilíbrio de classes	14
2.3.3	Generalização dos modelos, <i>overfitting</i> , dados de treino e dados de teste	14
2.4	Trabalho relacionado	14
<b>3</b>	<b>Metodologia</b>	<b>17</b>
3.1	Visão Geral	17
3.2	Dados	18
3.2.1	Dados de observação da Terra por satélite	18
3.2.2	Origem dos dados	19
3.2.3	Dados multi-temporais	19
3.2.4	Formato e tratamento inicial dos dados	20

3.2.5	Dimensão dos dados . . . . .	22
3.3	Classificação com floresta aleatória . . . . .	22
3.3.1	O classificador: <i>RandomForestClassifier</i> . . . . .	23
3.3.2	<i>Overfitting</i> : validação cruzada estratificada repetida . . . . .	24
3.3.3	Desequilíbrio das classes . . . . .	24
3.4	Classificação com RNC . . . . .	25
3.4.1	Pré-processamento dos dados . . . . .	25
3.4.2	Classificação multi-classe . . . . .	30
3.4.3	Classificação binária com um classificador RNC por categoria . . . . .	31
3.4.4	Escolha de parâmetros: <i>GridSearchCV</i> . . . . .	31
<b>4</b>	<b>Apresentação e discussão dos resultados</b>	<b>33</b>
4.1	Classificação binária com floresta aleatória . . . . .	33
4.2	Classificação multi-classe com RNC . . . . .	35
4.3	Classificação binária com RNC . . . . .	38
4.4	Análise e discussão dos resultados . . . . .	40
<b>5</b>	<b>Conclusões e trabalho futuro</b>	<b>41</b>
	<b>Referências</b>	<b>43</b>

# Lista de Figuras

2.1	Spectral signatures of different Earth features within the visible light spectrum. . . . .	6
2.2	Funcionamento do <i>stride</i> numa RNC. . . . .	11
2.3	Funcionamento do <i>padding</i> numa RNC. . . . .	11
2.4	Exemplo do resultado da aplicação de <i>GridSearchCV</i> para a RNC multi-classe. . . . .	12
2.5	Matriz de Confusão. . . . .	13
3.1	Classificação de EVU. Visão geral da metodologia. . . . .	18
3.2	Categorias a utilizar para a classificação do solo. . . . .	19
3.3	Esquema do processo associado ao classificador com floresta aleatória. . . . .	23
3.4	Exemplo de imagem 10x10 com pontos em falta. . . . .	27
3.5	Processo que incrementa o número de instâncias para entrada da rede convolucional. . .	27
3.6	Processo para adicionar linhas e/ou colunas aos parques de forma a que a largura e altura sejam múltiplos de 5. . . . .	28
3.7	Esquema do processo associado ao classificador com RNC. . . . .	30
3.8	RNC multi-classe: Definição da arquitetura do modelo. . . . .	31
3.9	RNC binária: Definição da arquitetura do modelo para cada categoria. . . . .	31
4.1	RNC multi-classe: Matriz de confusão com os valores reais vs valores previstos, para cada categoria. . . . .	36
4.2	RNC multi-classe: Matriz de confusão normalizada com os valores previstos divididos pelo valor real total de cada categoria. . . . .	37
4.3	RNC multi-classe: Matriz de confusão normalizada com os valores previstos divididos pelo valor previsto total de cada categoria. . . . .	38
4.4	RNC Binária: Matrizes de confusão da categoria <i>Coniferous</i> . . . . .	39



# Lista de Tabelas

3.1	Dimensão dos dados. . . . .	22
3.2	Desequilíbrio das categorias. . . . .	24
3.3	Resultado da aplicação de técnicas de correção de desequilíbrio de classes à categoria <i>Allotment</i> . . . . .	25
3.4	Parques por cidade. . . . .	26
3.5	Entrada da RNC-Classificação binária-Imagens por categoria. . . . .	29
3.6	Resultados do <i>Grid Search</i> para a RNC binária. . . . .	32
4.1	Resultados da classificação binária com floresta aleatória. . . . .	34
4.2	Classificação binária - <i>SmoteTomek</i> para a categoria <i>Coniferous</i> . . . . .	34
4.3	Classificação multi-classe-Imagens por categoria. . . . .	35
4.4	Resultados da classificação multi-classe com RNC. . . . .	35
4.5	Resultados da classificação binária com rede neuronal convolucional. . . . .	39





# Lista de abreviaturas

<b>ANN</b>	Artificial Neural Network
<b>CNN</b>	Convolutional Neural Network
<b>CV</b>	Cross Validation
<b>DL</b>	Deep Learning
<b>DM</b>	Data Mining
<b>DT</b>	Decision Trees
<b>FCUL</b>	Faculdade de Ciências da Universidade de Lisboa
<b>EVU</b>	Espaços Verdes Urbanos
<b>ML</b>	Machine Learning
<b>NN</b>	Neural Network
<b>ONU</b>	Organização das Nações Unidas
<b>RF</b>	Random Forest
<b>RNC</b>	Redes Neurais Convolucionais
<b>RNR</b>	Redes Neurais Recorrentes
<b>SVM</b>	Support Vector Machines
<b>UGS</b>	Urban Green Spaces



# Capítulo 1

## Introdução

### 1.1 Enquadramento e motivação

De acordo com o relatório da ONU [26], 55% da população mundial vive atualmente em cidades e as projeções apontam para que em 2050 essa percentagem atinja os 68%. A integração e adequação de espaços verdes nas cidades é um factor determinante na qualidade de vida dos seus habitantes. Os EVU são úteis para a produção de oxigénio, absorção de dióxido de carbono, regularização das temperaturas, controlo de humidade, controlo de raios solares, proteção contra o vento, chuva e granizo, proteção contra a erosão e absorção de águas pluviais, permitem separar fisicamente a circulação de automóveis e peões, reduzindo o ruído, filtrando os gases tóxicos e reduzindo o encandeamento. Além dos contributos para o bem estar das populações já referidos, constituindo espaços de lazer, proporcionam conforto psicológico, permitem uma ligação à natureza e contribuem para a biodiversidade.

A inventariação detalhada e precisa dos EVU, é fundamental para o planeamento urbano e para o desenvolvimento sustentável das cidades. O Atlas urbano da União Europeia [6] lançado em 2009 pela Comissão Europeia com o apoio da Agência Espacial Europeia e da Agência Europeia do Ambiente, disponibiliza um conjunto de mapas digitais de alta resolução para 305 áreas urbanas com mais de 100 000 habitantes dos 27 Estados-membros da União Europeia, com dados atualizados e precisos sobre a ocupação dos solos, obtidos através de imagens de satélite.

Os mapas existentes não detalham satisfatoriamente os EVU, porque são constituídos por áreas reduzidas mas apresentam alguma complexidade e muita diversidade, faltando informações importantes como a estrutura da vegetação e dos habitats aquáticos.

Assim, a motivação desta dissertação é contribuir para uma melhoria dos mapas existentes propondo algoritmos de aprendizagem automática aplicados a dados obtidos de imagens de satélite para classificar detalhadamente a estrutura da vegetação e dos habitats aquáticos nas áreas que representam os EVU.

## 1.2 Objectivos e resultados esperados

O objetivo desta dissertação é demonstrar a eficácia de algoritmos de aprendizagem automática, com particular ênfase em aprendizagem profunda, nomeadamente RNC, para a classificação do uso e cobertura de solo através de dados obtidos de imagens de satélite dos EVU, usando em simultâneo três diferentes tipos de variáveis explicativas, cores do espectro visível, valores de refletância e indicadores calculados a partir destes valores.

Espera-se que as métricas resultantes da avaliação dos modelos propostos, aplicados aos dados pré-classificados disponíveis, se situem em valores considerados adequados ao objetivo a que nos propusemos, de forma a poderem ser utilizados para uma correta classificação de novos dados, provenientes de outros EVU.

## 1.3 Contribuições

Para a prossecução dos objetivos descritos é proposta nesta dissertação a aplicação de 3 métodos de aprendizagem automática ao conjunto de dados constituído pelas coordenadas dos pontos geográficos, os valores de reflectância, indicadores calculados a partir daqueles e as percentagens atribuídas a cada categoria definida (*Allotment gardens*, *Artificial surfaces*, *Ornamental bed*, *Bare-lands*, *Grassland*, *Water surfaces*, *Riparian vegetation*, *Shrubs*, *Broad deciduous*, *Broad ever* e *Coniferous*):

- Classificação binária com um *ensemble* de árvores de decisão, floresta aleatória, aplicando um classificador a uma réplica do conjunto de dados para cada categoria.
- Classificação multi-classe com RNC aplicada a um conjunto de matrizes que representam imagens construídas através da transformação dos dados originais dos EVU de forma a que cada imagem corresponda a um espaço quadrado de 100 x 100 metros. A construção destas imagens foi efetuada sobrepondo pontos para incrementar o seu número e diversidade, já que a categoria atribuída a cada imagem, corresponde à mais representada no conjunto de pontos que a constituem. A entrada do modelo, no que corresponde à classificação da imagem, é constituída por 11 colunas correspondentes às categorias e para cada instância apenas uma das colunas é preenchida com 1, as restantes com 0 indicando, respetivamente, a ocorrência ou não ocorrência da categoria na imagem.
- Classificação Binária com RNC. Neste método, foi criada uma réplica para cada categoria do conjunto de imagens obtidas da mesma forma que no ponto anterior mas em que a entrada correspondente à classificação da imagem é constituído apenas por uma coluna preenchida com 1 ou 0 indicando, respetivamente, a ocorrência ou não ocorrência da categoria na imagem.

Os resultados obtidos nos dois modelos de redes convolucionais sugerem que a sua aplicação a dados disponíveis a partir de imagens de satélite pode constituir um contributo importante para a classificação da cobertura do solo dos EVU permitindo a sua adequada inventariação.

## 1.4 Estrutura do documento

A estrutura desta dissertação é constituída por:

- **Capítulo 2 Enquadramento e trabalho relacionado**, onde são descritos o problema em estudo, os conceitos de aprendizagem automática associados à tarefa de classificação a que nos propomos e indicação de trabalho anterior relacionado com o tema.
- **Capítulo 3 Metodologia** com a descrição dos dados, sua origem e dimensões, o tratamento efetuado para a sua utilização como entrada dos algoritmos de aprendizagem automática utilizados para a classificação, a descrição destes, bem como de métodos de otimização e regularização a eles associados.
- No **Capítulo 4 Apresentação e discussão dos resultados** são apresentados os resultados detalhados obtidos com cada um dos algoritmos, a sua análise e discussão.
- Finalmente, o **Capítulo 5 Conclusões e trabalho futuro**, apresenta as conclusões que podemos retirar dos resultados obtidos, melhorias possíveis e sugestões consideradas pertinentes para utilização ou adaptação do trabalho aqui desenvolvido, ao mesmo contexto ou a outro semelhante.



## Capítulo 2

# Enquadramento e trabalho relacionado

Este capítulo é composto por 3 secções. A primeira, inclui a descrição do problema, que consiste em classificar a composição dos EVU, as categorias definidas para essa classificação, a origem dos dados a utilizar, imagens de satélite, a sua origem e alguns conceitos associados a deteção remota. Na segunda secção, são apresentados os algoritmos mais populares para classificação em aprendizagem automática, sua caracterização, vantagens e desvantagens. São ainda descritos alguns conceitos e métodos utilizados nos modelos propostos. Na terceira secção, são abordadas a avaliação e seleção dos modelos, métricas utilizadas, problemas relacionados com estas tarefas e métodos para a sua resolução. Finalmente, na quarta secção são referidos alguns trabalhos relacionados com o problema em estudo nesta dissertação e considerados importantes para o seu enquadramento.

### 2.1 Descrição do problema

O desafio que assumimos é detalhar de forma correta e precisa a composição dos EVU, através do tipo de cobertura de ocupação do solo, *Allotment, Artificial, Ornamental bed, Bare-lands, Grassland, Water surfaces, Riparian vegetation, Shrubs, Broad deciduous, Broad ever e Coniferous*, utilizando como variáveis preditivas, as cores do espectro visível, os valores da refletância das imagens multiespectrais, ambas obtidas através do Satélite Sentinel-2 por deteção remota e indicadores calculados a partir dos valores da refletância (índices de solo, densidade urbana, água, vegetação e variáveis biofísicas).

A deteção remota é uma forma de obter informações acerca dos objetos, recolhendo e analisando dados sem que os instrumentos usados para os recolher estejam em contacto directo com o objecto [10]. Os sensores remotos são sistemas que registam o fluxo de energia eletromagnética refletida por objetos distantes. Existem dois tipos de sensores remotos, os **ativos**, que emitem eles próprios a radiação em direção aos objetos que pretendem detetar e os **passivos**, que detetam a radiação natural que todos os objetos com temperaturas acima do zero absoluto emitem ou a radiação emitida pela luz solar e refletida

pelos objetos.

A refletância é a *relação entre a intensidade de luz que incide numa superfície e a luminosidade que é por ela refletida* [16]. É esta a variável medida diretamente pelos instrumentos de deteção remota (sensores) e corresponde a um rácio entre a quantidade de luz refletida pelo objeto e a quantidade de luz que ele recebe.

Diferentes tipos de superfície, como água, solo nu e vegetação, refletem a radiação de maneira diferente em vários canais (ver Figura 2.1). A radiação refletida em função do comprimento de onda é chamada de assinatura espectral da superfície [11].

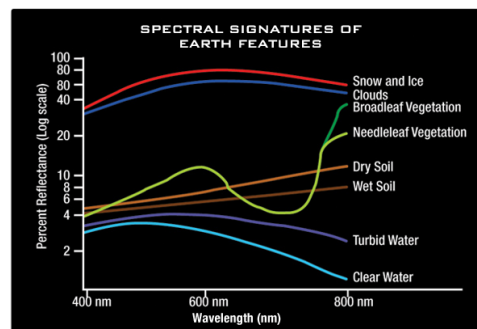


Figura 2.1: Spectral signatures of different Earth features within the visible light spectrum.

Credit: Jeannie Allen<sup>1</sup>

## 2.2 Algoritmos de aprendizagem automática para classificação

A Classificação de dados é um processo de aprendizagem automática constituído por 2 passos:

**Aprendizagem:** Corresponde à fase de construção do modelo;

**Classificação:** quando usamos o modelo construído para prever qual a classe a que pertence o objeto que pretendemos identificar [13]. A classificação pode ser binária, quando os objetos a classificar apenas podem pertencer a 2 classes ou multi-classe, quando podemos atribuir uma de várias classes ao objeto, ou ainda *multi-label* em que podemos atribuir várias classes ao mesmo objeto.

Nos pontos seguintes apresentamos os algoritmos de classificação mais populares, respetivas características, vantagens e desvantagens de cada um deles.

<sup>1</sup><https://towardsdatascience.com/satellite-imagery-analysis-using-python-9f389569862c>



### 2.2.1 SVM

Funciona classificando os dados em diferentes classes, localizando uma linha (hiper-plano) que separa o conjunto de dados de treino em classes. Como existem muitos hiper-planos lineares, o algoritmo SVM tenta maximizar a distância entre as várias classes envolvidas a que se dá o nome de maximização da margem. Se a linha que maximiza a distância entre as classes for identificada, aumentará a probabilidade de generalização para novos dados.

**Vantagens:** funciona relativamente bem quando há uma clara margem de separação entre as classes; é mais eficaz em espaços de alta dimensão; é eficaz nos casos em que o número de dimensões é maior que o número de amostras; é relativamente eficiente em consumo de memória [17].

**Desvantagens:** não é adequado para grandes conjuntos de dados; não funciona muito bem quando o conjunto de dados tem mais ruído, ou seja, as classes de destino estão sobrepostas; nos casos em que o número de recursos para cada ponto de dados exceder o número de amostras de dados de treino, a aprendizagem será insuficiente; como o SVM funciona colocando pontos de dados acima e abaixo do hiper-plano, não há explicação probabilística para a classificação. [17]

### 2.2.2 Árvores de decisão

Uma árvore de decisão é uma representação gráfica que utiliza a metodologia de ramificação para exemplificar todos os resultados possíveis de uma decisão, com base em determinadas condições. Numa árvore de decisão, o nó interno representa um teste no atributo, cada ramo da árvore representa o resultado do teste e o nó folha representa um rótulo de classe específico, isto é, a decisão tomada após o cálculo de todos os atributos. As regras de classificação são representadas pelo caminho da raiz ao nó da folha.[13]

**Vantagens:** Comparando com outros algoritmos, as árvores de decisão requerem menos esforço para a preparação de dados durante o pré-processamento; uma árvore de decisão não requer normalização de dados; uma árvore de decisão não requer dimensionamento de dados; valores ausentes nos dados não afetam significativamente o processo de construção da árvore de decisão; um modelo de árvores de decisão é muito intuitivo e fácil de explicar para equipas técnicas e partes interessadas.

**Desvantagens:** Uma pequena alteração nos dados pode causar uma grande alteração na estrutura da árvore de decisão, causando instabilidade; por vezes, o cálculo pode ser muito mais complexo em comparação com outros algoritmos; a árvore de decisão geralmente envolve mais tempo para treinar o modelo; o treino em árvore de decisão é relativamente caro, pois é mais complexo e demorado; o algoritmo da árvore de decisão é inadequado para aplicar a regressão e prever valores contínuos.

### 2.2.3 Redes neuronais artificiais

Uma rede neuronal é um processador paralelamente distribuído, massivamente, composto de unidades de processamento simples que tem uma propensão natural para armazenar conhecimento experimental e disponibilizá-lo para uso [14]. Assemelha-se ao cérebro em dois aspetos:

1. O conhecimento é adquirido pela rede a partir do seu ambiente, através de um processo de aprendizagem.
2. Os pontos fortes da conexão interneurónio, conhecidos como pesos sinápticos, são usados para armazenar o conhecimento adquirido.

A rede neuronal artificial é um sistema de computação projetado para replicar a maneira como o cérebro humano funciona, em geral, inspirado na estrutura das redes neuronais biológicas do cérebro humano. Existe um neurónio de entrada que atua como um interface para todos os outros neurónios passarem pela entrada e existe um neurónio de saída que aceita todas as saídas de diferentes neurónios[21].

**Vantagens:** As redes neuronais têm a capacidade de aprender por si mesmas e produzir um output que não se limita ao conhecimento que obtiveram no processo de aprendizagem; o input é armazenado internamente e não em bases de dados externas por isso não são afetadas por eventuais perdas de dados; podem aprender com os exemplos e aplicá-los quando ocorrer um evento semelhante, tornando-as capazes de trabalhar com eventos em tempo real; mesmo se um neurónio não responder ou faltar uma informação, a rede poderá detetar a falha e ainda assim produzir o output; podem executar várias tarefas em paralelo sem afetar o desempenho do sistema [21].

**Desvantagens:** Os algoritmos de redes neuronais funcionam como uma espécie de caixa negra não permitindo perceber porque é tomada uma determinada decisão de classificação; o tempo de desenvolvimento de um modelo baseado neste método pode ser muito superior comparativamente a outros algoritmos mais simples; geralmente exigem muito mais dados do que outros algoritmos; são computacionalmente mais caros [8].

### 2.2.4 Redes neuronais profundas

A aprendizagem profunda ou *deep learning* é uma evolução das rede reuronais e distingue-se destas por funcionar utilizando múltiplas camadas ocultas em cascata. Existem tipos diferentes de redes neuronais profundas e cada um deles possui vantagens e desvantagens, dependendo do uso [24]:

- **Redes neuronais convolucionais:** Contêm cinco tipos de camadas, de entradas, de convolução, de agrupamento, as completamente conetadas e as de saída. Cada camada tem um propósito específico, como seja, de resumo, conexão ou ativação. As redes convolucionais popularizaram a

classificação de imagens e a deteção de objetos, mas também têm vindo a ser aplicadas em outras áreas como previsão e processamento de linguagem natural.

- **Redes neuronais recorrentes:** Usam informações sequenciais, como dados de registo de data e hora de um sensor ou uma frase dita. Essas informações são compostas por uma sequência de termos. Ao contrário das redes neuronais tradicionais, as entradas de uma rede neuronal recorrente não são independentes umas das outras, e os resultados para cada elemento dependem da computação dos elementos precedentes. Redes recorrentes são utilizadas na previsão e aplicação de séries temporais, análise de sentimento e outras aplicações de texto.
- **Redes neuronais *feedforward*:** Nas quais cada camada é completamente conetada à camada seguinte. A informação é entregue de maneira antecipada de uma camada à seguinte seguindo sempre em frente. Não há *loops* de *feedback*.
- **Redes neuronais *autoencoder*:** São utilizadas para criar abstrações chamadas *encoders*, criados a partir de um conjunto estipulado de entradas. Apesar de similares às redes neuronais mais tradicionais, *autoencoders* procuram modelar as entradas por si só e, portanto, o método é considerado não supervisionado. A premissa dos *autoencoders* é diminuir a sensibilidade ao que é irrelevante e aumentar ao que é. à medida que as camadas vão sendo adicionadas, outras abstrações são formuladas em camadas mais altas (camadas mais próximas do ponto onde uma camada descodificadora é introduzida). Essas abstrações podem, então, ser usadas por classificadores lineares ou não lineares.

### 2.2.5 Métodos *ensemble*

Os modelos *ensemble* fazem previsões com base em vários modelos diferentes. Estes modelos, ao combinarem modelos individuais, tendem a ser mais flexíveis (menos enviesamento) e menos sensíveis à variabilidade dos dados [19]. Existem diversos tipos de *ensemble*, sendo os mais populares o *bagging* e o *boosting*:

- ***Bagging*:** Um grupo de modelos individuais é treinado em paralelo. Cada modelo é treinado utilizando um sub-conjunto aleatório de dados.
- ***Boosting*:** O grupo de modelos é treinado sequencialmente. Cada modelo individual aprende com os erros do modelo anterior, através da alteração dos pesos das instâncias mal classificadas.

### 2.2.6 Floresta aleatória

É um modelo ensemble que consiste na utilização de várias árvores de decisão, em que para cada uma delas é escolhido aleatoriamente um subconjunto de dados. A classificação final é obtida por votação, ou seja, a classificação que obtiver o maior número de resultados entre todas as árvores é a escolhida. Este algoritmo utiliza geralmente o método de *bagging*, embora também possa ser usado outro método chamado *forest-RC* em que, em vez de vários subconjuntos aleatórios de dados, são criados novos atributos através de uma combinação linear de atributos existentes [13], [19].

### 2.2.7 Extreme gradient boosting

O *eXtreme Gradient Boosting* (XgBoost) é uma conhecida e eficiente implementação de código aberto do algoritmo baseado em árvores com aumento de gradiente. O aumento de gradiente é um algoritmo de aprendizagem supervisionada que tenta prever com precisão uma variável de destino. Para isso, combina um grupo de estimativas de um conjunto de modelos mais simples e mais fracos. O *XgBoost* tem excelente desempenho em competições de aprendizagem automática porque lida de maneira robusta com uma variedade de tipos de dados, relacionamentos e distribuições e por causa do grande número de hiperparâmetros que podem ser aperfeiçoados e ajustados para um cenário mais apropriado. Essa flexibilidade faz do *XgBoost* uma escolha consistente para problemas de regressão, classificação (binária e multi-classe) e pontuação [19], [2].

### 2.2.8 Componentes de aprendizagem profunda: *Stride* e *Padding*

Pela importância que representam e para enquadrar a utilização de dois processos semelhantes adotados neste trabalho, descrito no capítulo **Metodologia**, referimos aqui dois componentes fundamentais dos algoritmos de aprendizagem profunda para classificação de imagens [28].

- **Stride**: é um parâmetro do filtro utilizado nas redes neuronais convolucionais para o tratamento de imagens que define quanto é que o filtro se desloca ao longo da imagem reduzindo a sua dimensão, como ilustra a Figura 2.2.
- **Padding**: é um processo de adicionar pixels brancos ou vazios a uma imagem para compensar a redução de tamanho e permitir que os píxeis que estão nas margens da imagem não fiquem sub-representados, como ilustra a Figura 2.3.

---

<sup>2</sup><https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/>

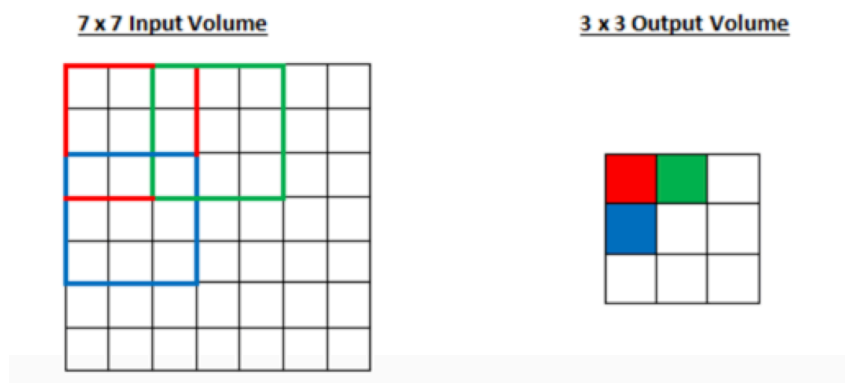


Figura 2.2: Funcionamento do *Stride* numa RNC.

Fonte: Adit Deshpande<sup>2</sup>

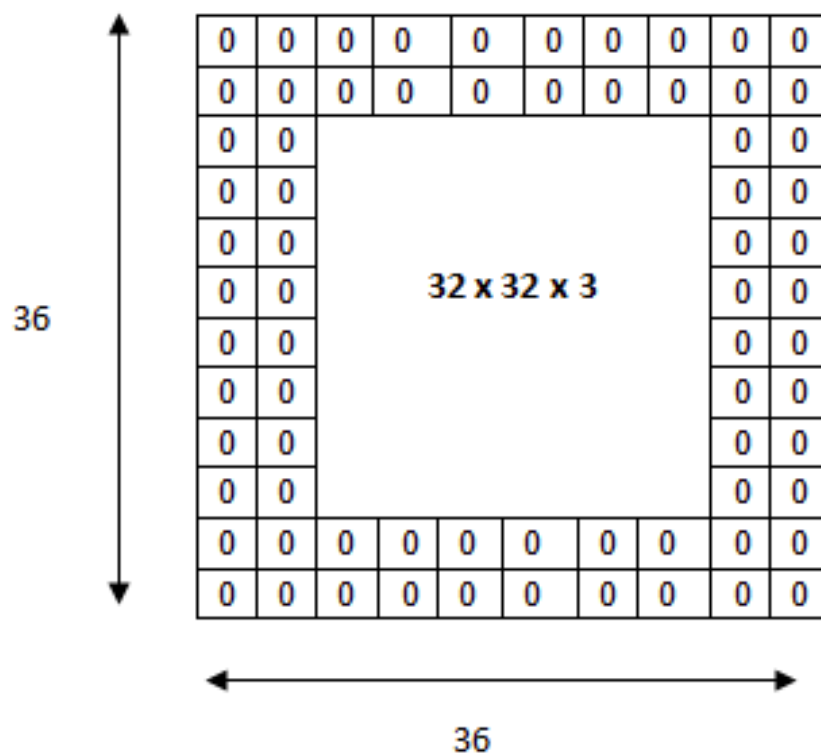


Figura 2.3: Funcionamento do *padding* numa RNC.

Fonte: Adit Deshpande<sup>3</sup>

<sup>3</sup><https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/>

### 2.2.9 Escolha de parâmetros: *GridSearchCV*

Os hiperparâmetros dos estimadores são argumentos incluídos na sua construção, o que significa que não são aprendidos durante o processo de aprendizagem, mas são decisivos para o desempenho do modelo. Escolher qual o valor ou a combinação de valores dos hiperparâmetros é uma tarefa dispendiosa, pois as possibilidades são muitas e não existem regras científicas para a sua escolha, pois variam consoante a arquitetura do modelo e os dados. Existem por isso métodos que facilitam esta escolha. É o caso do *GridSearchCV* da biblioteca *scikit-learn* [22], que utilizámos para a escolha de *batch\_size* e *epochs*, hiperparâmetros da RNC. A *GridSearchCV* implementa a API do estimador e ajusta-a ao conjunto de dados, avalia todas as combinações possíveis de valores dos parâmetros, utilizando validação cruzada e retorna os resultados indicando qual a melhor combinação, como exemplificado na Figura 2.4.

```
Best: 0.671817 using {'batch_size': 30, 'epochs': 50}
0.649673 (0.237089) with: {'batch_size': 10, 'epochs': 10}
0.660455 (0.247096) with: {'batch_size': 10, 'epochs': 25}
0.656258 (0.232080) with: {'batch_size': 10, 'epochs': 50}
0.635982 (0.246733) with: {'batch_size': 10, 'epochs': 100}
0.645356 (0.249486) with: {'batch_size': 20, 'epochs': 10}
0.659596 (0.231659) with: {'batch_size': 20, 'epochs': 25}
0.649367 (0.249090) with: {'batch_size': 20, 'epochs': 50}
0.663886 (0.252225) with: {'batch_size': 20, 'epochs': 100}
0.638617 (0.234148) with: {'batch_size': 30, 'epochs': 10}
0.652583 (0.241738) with: {'batch_size': 30, 'epochs': 25}
0.671817 (0.246307) with: {'batch_size': 30, 'epochs': 50}
0.636534 (0.245153) with: {'batch_size': 30, 'epochs': 100}
0.651756 (0.247148) with: {'batch_size': 40, 'epochs': 10}
0.656259 (0.246812) with: {'batch_size': 40, 'epochs': 25}
0.663026 (0.239695) with: {'batch_size': 40, 'epochs': 50}
0.652062 (0.248410) with: {'batch_size': 40, 'epochs': 100}
0.639290 (0.244033) with: {'batch_size': 50, 'epochs': 10}
0.662476 (0.240109) with: {'batch_size': 50, 'epochs': 25}
0.652828 (0.238640) with: {'batch_size': 50, 'epochs': 50}
0.651664 (0.246279) with: {'batch_size': 50, 'epochs': 100}
```

Figura 2.4: Exemplo do resultado da aplicação de *GridSearchCV* para a RNC multi-classe.

## 2.3 Avaliação e seleção dos modelos

A utilização de aprendizagem automática para classificação consiste, basicamente, em treinar um algoritmo com dados para os quais sabemos a classe a que pertencem (pré-classificados), de forma a obter um modelo que nos permita posteriormente efetuar predições a partir de novos dados não classificados. A seleção do modelo é então baseada no seu desempenho que pode ser medido através de métricas que veremos a seguir. Nesta secção vamos também referir os problemas relacionados com o processo de

classificação e técnicas para os resolver ou minimizar.

### 2.3.1 Métricas para avaliar o desempenho dos classificadores

Para referirmos as métricas usadas para avaliar os modelos é importante ter presente os conceitos de casos positivos e casos negativos. Dizemos que uma instância é um caso positivo se pertence à classe de interesse e é negativo se pertence a outra classe qualquer. Se tivermos um problema de classificação binária, a instância ou objeto apenas pode pertencer ou não pertencer à classe ou categoria em estudo. Posto isto, convém agora referir que o processo de avaliação do modelo consiste em medir de que forma é que ele classificou as instâncias. Assim, se o modelo classificou corretamente um exemplo, temos um verdadeiro positivo (TP - *true positive* do inglês), caso ele pertença à classe de interesse ou verdadeiro negativo (TN - *true negative* do inglês) caso não pertença. Se o exemplo pertence à classe de interesse e o modelo o classificou como não pertencendo, temos um Falso Negativo (FN), caso o exemplo não pertença à classe de interesse e o modelo o classificou positivamente estamos na presença de um Falso Positivo (FP).

A Matriz de Confusão é uma técnica muito útil para analisar até que ponto o classificador consegue reconhecer corretamente instâncias de diferentes classes. Observando a Figura 2.5, podemos deduzir que se o nosso classificador fosse perfeito, apenas a diagonal da matriz teria valores diferentes de 0 e a soma dos seus elementos seria igual ao número de instâncias, ou seja, não haveria Falsos Positivos nem Falsos Negativos, todas as instâncias estariam corretamente classificadas.

		Classe Prevista		Total
		Sim	Não	
Classe Real	Sim	TP	FN	P
	Não	FP	TN	N
Total		P'	N'	P+N

Figura 2.5: Matriz de Confusão.

A partir dos conceitos descritos podemos chegar às métricas mais comuns para a avaliação do desempenho dos modelos:

**Accuracy:** mede a proporção de instâncias corretamente classificadas do total de instâncias.

$$\frac{TP + TN}{P + N}$$

**Precision:** mede a proporção de instâncias realmente positivas das instâncias classificadas como positivas.

$$\frac{TP}{TP + FP}$$

**Recall:** mede a proporção de instâncias classificadas como positivas do total de instâncias realmente positivas.

$$\frac{TP}{P}$$

**Specificity:** mede a proporção de instâncias classificadas como negativas do total de instâncias realmente negativas.

$$\frac{TN}{N}$$

### 2.3.2 O problema do desequilíbrio de classes

Em muitas situações de classificação de objetos é comum termos um desequilíbrio de classes. Isto acontece quando a maioria das instâncias pertence a uma classe, em classificação binária podemos dizer que são na maioria negativas ou ao contrário, maioritariamente positivas. As 2 últimas métricas referidas no ponto anterior podem ser usadas para medir de que forma o modelo é capaz de reconhecer as instâncias positivas (*recall*, ou sensibilidade) ou negativas (especificidade).

Para combater este problema existem métodos que permitem adicionar (*oversampling*) ou retirar (*undersampling*) instâncias, de forma a equilibrar as classes. São exemplo destes métodos, respetivamente os algoritmos *SMOTE*[5] e *Tomek links*[25]. Há ainda combinações entre os dois métodos, como é o caso do *imblearn.combine.SMOTETomek*, disponível na *imbalanced-learn API* da *scikit-learn*.

### 2.3.3 Generalização dos modelos, *overfitting*, dados de treino e dados de teste

Um problema recorrente na aplicação de métodos de aprendizagem automática para classificação é obtermos excelentes resultados durante o treino mas depois o modelo ser inútil na presença de novas instâncias. Por vezes, o modelo não é generalizável devido a um fenómeno conhecido por *overfitting*.

A solução para este problema é utilizar dois conjuntos de dados completamente separados, um para treinar o modelo e outro para a sua avaliação. Tipicamente, existem 3 métodos cada um com uma diferente estratégia para abordar esta solução, *Holdout*, *Cross-validation* e *Bootstrap*.

Por ser um método que funciona particularmente bem em aprendizagem supervisionada[12], que é o caso presente, focar-nos-emos no método *k-fold cross-validation*. Este método consiste em dividir o conjunto inicial de dados em *k* subconjuntos ou *folds* com aproximadamente o mesmo tamanho. Em cada uma das *k* iterações o modelo aprende em *k-1 folds* e é testado com aquele que ficou de fora. No final, considera-se a média das métricas obtidas em cada uma das *k* iterações. Empiricamente, é recomendado o uso de 10-*folds*, devido ao baixo viés e variância que apresenta.

## 2.4 Trabalho relacionado

Nos anos mais recentes, diversos e com variados objetivos têm sido os trabalhos realizados sobre dados de observação da terra, utilizando métodos de Aprendizagem Automática para classificar cobertura



e uso do solo. Aqui se descrevem alguns dos mais relevantes.

Castelluccio et al. [4], utilizaram RNC com duas diferentes arquiteturas, CaffèNet e GoogleLeNet e três distintos tipos de aprendizagem, *from scratch*, *fine-tuning* e *feature vector*, para classificar diversos tipos de uso da terra. Este trabalho incidiu sobre 2 *datasets*, o *UC Merced* [27], que é um conhecido *dataset* público com 21 classes de uso da terra com 100 imagens de detecção remota de várias regiões dos EUA e um conjunto de imagens de plantações de café do estado de Minas Gerais no Brasil, com 3 classes, *coffee*, *non-coffee* e *mixed* [23]. As conclusões do estudo revelam uma clara superioridade do método aprendizagem profunda, sobre os melhores métodos de referência.

Maggiore et al. [20] propuseram uma *Fully Convolutional Neural Network*, para reduzir o tempo de execução do processo de aprendizagem para classificação de imagens. Este tipo de rede neuronal é proposto como alternativa às abordagens tradicionais e a um outro tipo designado por *Patch-based*. Os resultados experimentais mostram que esta abordagem supera as anteriores tanto na precisão como no tempo computacional.

Kussul et al. [18] utilizaram imagens multitemporais dos satélites LandSat-8 e Sentinel-1A, para identificar 11 classes de ocupação do solo agrícola e tipos de cultivo: água, floresta, pastagem, terra limpa, trigo de inverno, colza, cereais de primavera, soja, milho, girassol e beterraba sacarina. Foram propostas redes neurais convolucionais, como alternativa aos métodos mais tradicionais de classificação de imagens, um *ensemble de MLP's (Multilayer Perceptrons)* e *Random Forest*. Os resultados evidenciaram um melhor desempenho da arquitetura constituída pelo *ensemble* de RNC relativamente às restantes.

Helber et al. [15] com o objetivo de melhorarem os mapas geográficos existentes, propuseram-se a classificar o uso e cobertura da terra em 30 diferentes países da União Europeia, apresentando um novo conjunto de dados baseado em imagens do satélite *Sentinel-2*. Este *dataset* consiste em 10 classes cobrindo 13 diferentes bandas espectrais num total de 27.000 imagens. Usaram redes neurais profundas com as arquiteturas GoogleNet e ResNet50 demonstrando a supremacia deste método de classificação sobre outras abordagens, tanto para o novo *dataset* como para dados de detecção remota previamente existentes. Concluiu-se também que os novos dados oferecem melhores resultados.

Abdi [1] aplicou vários métodos de Aprendizagem Automática (SVM, *Random Forest*, Xgboost, Aprendizagem Profunda) para classificação do uso e cobertura da terra, neste caso da região de Uppsala no centro-sul da Suécia *Boreal Landscapes*. Foram utilizadas imagens multi-temporais (abril, maio, junho e outubro de 2017) com 10 bandas espectrais obtidas no âmbito do programa *Copernicus* do satélite *Sentinel-2*. A classificação compreende um total de 25 classes organizadas em 8 classes principais, (1) Floresta decídua: compreendendo bétula, álamo, amieiro, faia, carvalho, olmo e freixo; (2) Floresta de coníferas: compreendendo pinheiro silvestre e abeto da Noruega; (3) Água: lagos, rios e canais; (4) Artificial: áreas urbanas, canteiros de obras e estradas; (5) Áreas húmidas: terras saturadas, incluindo pântanos e paúis; (6) agricultura; (7) Cortes nítidos dentro e fora de áreas húmidas; e (8) terrenos abertos com e sem vegetação. As sete classes restantes foram descartadas devido à presença insignificante na área de estudo. Nas conclusões deste estudo são destacados alguns pontos que consideramos serem muito relevantes no contexto do estudo que vamos propor, a saber: (1) O instrumento multi-espectral *Sentinel-2*

é único entre os satélites de observação da terra atualmente em operação, devido às suas 3 bandas de *red edge* que podem capturar o conteúdo de clorofila da planta e à sua resolução espacial média-alta de 10 metros; (2) Dos 4 algoritmos de aprendizagem automática ensaiados, dois deles, SVM e floresta aleatória são mais populares na comunidade de *Remote Sensing* (RS) e os outros dois, *Xgboost* e aprendizagem profunda na comunidade de ciência de dados embora estejam a ganhar popularidade também na comunidade de RS. Os 4 algoritmos produziram classificações semelhantes com uma precisão entre 0,733 e 0,758, no entanto a aprendizagem profunda teve o desempenho mais baixo, provavelmente devido a um grande número de classes e à saturação da função de ativação; (3) O facto de um quarto das principais bandas deste estudo serem *red edge* sugere que os índices de vegetação espectral, que são frequentemente utilizados em classificação de uso e cobertura do solo, para distinguir entre tipos de vegetação podem perder relevância no contexto de classificação de regiões dominadas por florestas; (4) Relativamente ao uso de imagens multi-temporais, as imagens de maio e julho foram predominantes no contributo para as classificações obtidas, o que permite questionar se o uso de imagens mono-temporais no período certo não será preferível ao uso de imagens multi-temporais.

# Capítulo 3

## Metodologia

Este capítulo está organizado em quatro secções. Na primeira, será apresentada uma visão geral da metodologia e quais os algoritmos de aprendizagem automática utilizados para a resolução do problema. Em seguida, são descritos os dados utilizados, sua origem, tipo, formato e tratamento inicial a que foram submetidos. Nas duas restantes secções, apresentamos em pormenor os algoritmos utilizados e o pré-processamento efetuado aos dados, quer para se adequarem a cada um dos 3 modelos, quer para lidar com problemas como sejam, o desequilíbrio das classes e o *overfitting*.

### 3.1 Visão Geral

Nesta secção, é apresentada uma descrição esquematizada do problema e das propostas para a sua solução.

A Figura 3.1 apresenta uma visão geral do problema e da metodologia a utilizar, quais as fontes de dados e as principais fases do trabalho. Para atingirmos o objetivo proposto, obtenção de um ou vários modelos (podemos, no limite, ter um modelo para cada classe) de classificação da vegetação e habitats aquáticos dos EVU iremos testar e avaliar alguns algoritmos de aprendizagem automática adequados a processos de classificação. O nosso foco serão os algoritmos de aprendizagem profunda, porque as características do nosso problema e a análise do trabalho relacionado sugerem essa abordagem. No entanto, vimos também nessa análise que os resultados de cada método variam em função do tipo de cobertura do solo a classificar e como não temos referência específica para a classificação em EVU, propomo-nos a testar o desempenho de vários algoritmos, nomeadamente floresta aleatória e aprendizagem profunda com RNC.

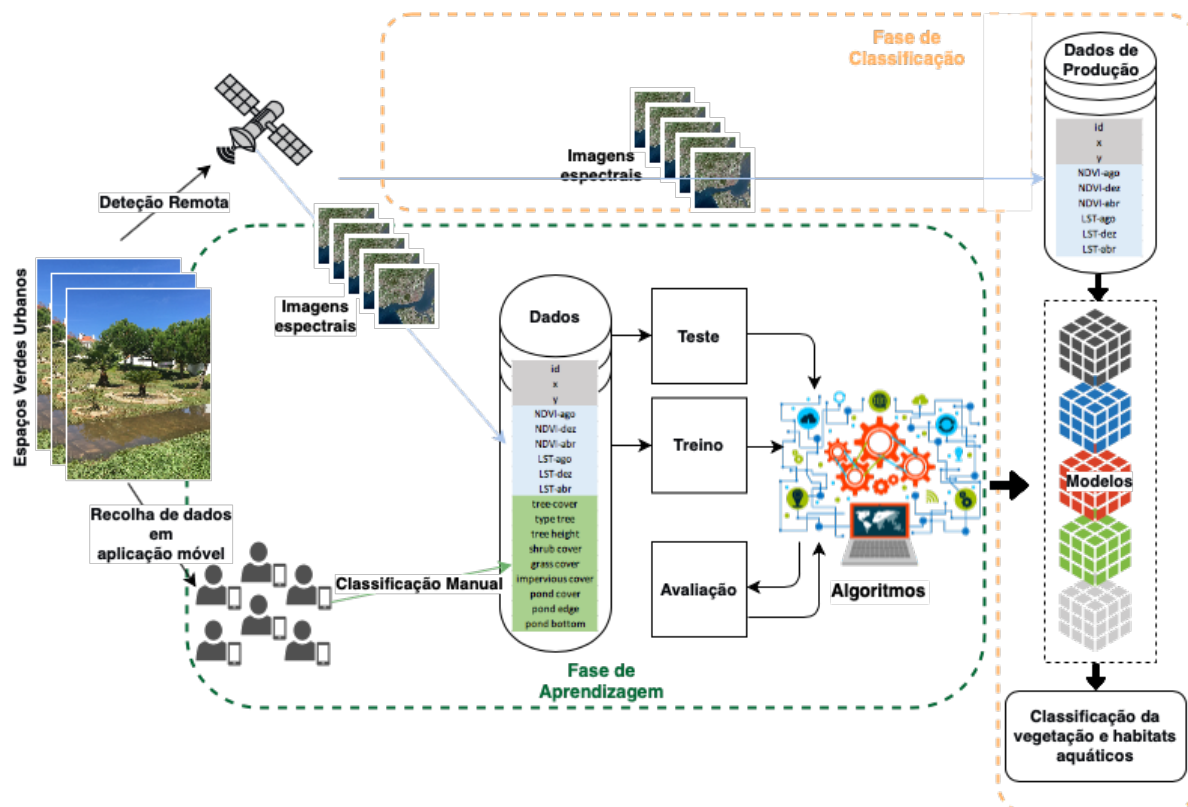


Figura 3.1: Classificação de EVU. Visão geral da metodologia.

## 3.2 Dados

Esta secção é constituída por uma descrição pormenorizada dos dados: tipo de dados, origem, formato, dimensão e tratamento inicial.

### 3.2.1 Dados de observação da Terra por satélite

Os dados a tratar correspondem a imagens com origem no satélite Sentinel-2, obtidos através do Copernicus, programa europeu para a Observação da Terra, criado pelo Regulamento 377/2014 do Parlamento Europeu e do Conselho Europeu. É uma parceria estabelecida entre a União Europeia, a Agência Espacial Europeia e os vários Estados-Membros [7]. São imagens multi-espectrais de ampla faixa e alta resolução, 13 bandas espectrais (quatro bandas a 10 metros, seis bandas a 20 metros e três bandas com resolução espacial de 60 metros). A missão do Sentinel-2 apoia aplicativos operacionais principalmente para serviços terrestres, incluindo a monitorização da vegetação, cobertura do solo e da água, bem como a observação de vias navegáveis interiores e áreas costeiras e complementa missões como SPOT ou Landsat[9].

Os conjuntos de dados a tratar contêm a identificação das imagens através das coordenadas geográ-

ficas no formato graus e minutos decimais (e.g., 38.721970°, -9.139337°), a referência temporal (mês, dia) e as variáveis explicativas que compreendem os valores das refletâncias de várias bandas espectrais e os valores dos índices de solo, de densidade urbana, de água, de vegetação e de variáveis biofísicas.

### 3.2.2 Origem dos dados

A existência de grandes quantidades de dados pré-classificados é determinante para o sucesso da classificação a efetuar. Nesse sentido, este trabalho está relacionado com um projeto a decorrer em paralelo que prevê o desenvolvimento de uma aplicação móvel, para registar localmente, por observação direta, as características dos tipos de cobertura do solo nos EVU ou mais concretamente classificar um espaço físico, em forma de quadrado (10m x 10m), de acordo com as categorias apresentadas na Figura 3.2).

Categorias (Tipo de cobertura do solo)			
Allotment	Bare-Lands	Riparian Vegetation	Broad Ever
Artificial	Grassland	Shrubs	Coniferous
Ornamental Bed	Water Surfaces	Broad Deciduous	

Figura 3.2: Categorias a utilizar para a classificação do solo.

Como os dados classificados desta forma não ficarão disponíveis em tempo útil, pois as datas de término dos dois projetos são coincidentes, vamos utilizar um conjunto de dados pré-classificados obtidos no âmbito do projeto Bioveins [3]. Este projeto é uma parceria entre várias Universidades Europeias e a *BiodivERsA - the network programming and funding research on biodiversity and ecosystem services across European countries and territories*. O conjunto de dados respeita a 220 EVU em 7 cidades europeias (Lisboa, Almada, Zurique, Paris, Antuérpia, Poznan e Tartu). A Universidade de Lisboa é uma das Universidades participantes e por isso os dados foram extraídos e preparados pelo Departamento de Biologia Vegetal da FCUL. Desta forma, podemos desenvolver os modelos a partir destes dados e fica em aberto a possibilidade de os refinar com dados alargados a outros espaços, obtidos posteriormente através da aplicação móvel que está atualmente em desenvolvimento.

### 3.2.3 Dados multi-temporais

A utilização de dados multi-temporais neste trabalho é justificada pelo facto de estarem a ser classificados elementos naturais que obviamente alteram significativamente ao longo do ano, em função das estações. Essas alterações implicam diferentes refletâncias a serem registadas pelos sensores e daí podem surgir distintas classificações da mesma vegetação, consoante o período do ano a que se referem os

dados, o que pode afetar o desempenho do modelo final.

### 3.2.4 Formato e tratamento inicial dos dados

Os dados para a aprendizagem foram disponibilizados através de 7 ficheiros *csv*, um por cada cidade contendo 3 conjuntos de colunas:

1. Identificação do ponto geográfico:

- **City**: Cada uma das 7 cidades, Almada, Antuérpia, Lisboa, Paris, Poznan, Tartu e Zurique.
- **Row\_Label**: N° sequencial no ficheiro.
- **X**: Longitude do centróide, no sistema de coordenadas projetado *ETRS\_1989\_LAEA - Lambert\_Azimuthal\_Equal\_Area*.
- **Y**: Latitude do centróide, no sistema de coordenadas projetado *ETRS\_1989\_LAEA - Lambert\_Azimuthal\_Equal\_Area*.

Cada elemento corresponde a um espaço de 10 metros por 10 metros, por isso X e Y representam o centróide desse espaço.

2. Classificação do ponto em termos de cobertura do solo efetuada a partir de observação direta dos espaços físicos:

- **Allotment gardens**
- **Artificial surfaces**
- **Bambu**
- **Bare-lands**
- **Flower beds**
- **Grassland**
- **Lakes and ponds**
- **Riparian vegetation**
- **Shrubs**
- **Rivers and streams**
- **Broad deciduous**
- **Broad ever**
- **Coniferous**

Durante a fase de pré-processamento dos dados, algumas das categorias foram agregadas, pois tinham uma representatividade muito baixa, ficando apenas 11 classes, como é visível na Tabela

3.1.

Através de observação direta no local, foi atribuído um valor em porcentagem a cada classe indicando a proporção da sua ocorrência no espaço respectivo, variando por isso entre 0 e 100%. Como a abordagem efetuada consistiu em classificar cada espaço em função da ocorrência ou não ocorrência de cada categoria, estes valores foram alterados para 1, quando diferentes de 0.

3. Características do ponto obtidas a partir da imagem de satélite:

- **Red:** Espectro visível, Gama 0-255.
- **Green:** Espectro visível, Gama 0-255.
- **Blue:** Espectro visível, Gama 0-255.
- **B02:** *Blue band* (Bandas espectrais de refletância).
- **B03:** *Green band* (Bandas espectrais de refletância).
- **B04:** *Red band* (Bandas espectrais de refletância).
- **B05:** *Red Edge 1* (Bandas espectrais de refletância).
- **B06:** *Red Edge 2* (Bandas espectrais de refletância).
- **B07:** *Red Edge 3* (Bandas espectrais de refletância).
- **B08:** *Near Infra Red* (Bandas espectrais de refletância).
- **B8A:** *Narrow Near Infra Red* (Bandas espectrais de refletância).
- **B11:** *SWIR-1* (Bandas espectrais de refletância).
- **B12:** *SWIR-2* (Bandas espectrais de refletância).
- **BI2:** *Second Brightness Index* (Índices de Solo).
- **CI:** *Colour Index* (Índices de Solo).
- **UI:** *Urban Index* (Índices de Densidade Urbana).
- **NDBI:** *Normalized Difference Build-Up Index* (Índices de Densidade Urbana).
- **NWDI:** *Normalized Difference Water Index* (Índices de Água).
- **MNWDI:** *Modified Normalized Difference Water Index* (Índices de Água).
- **NDVI:** *Normalized Difference Vegetation Index* (Índices de Vegetação).
- **SAVI:** *Soil Adjusted Vegetation Index* (Índices de Vegetação).
- **GNDVI:** *Green Normalized Difference Vegetation Index* (Índices de Vegetação).
- **FAPAR:** *Fraction of Absorbed Photosynthetically Active Radiation* (Variáveis Biofísicas).
- **FCOVER:** *Fraction of vegetation cover* (Variáveis Biofísicas).
- **LAI:** *Leaf Area Index* (Variáveis Biofísicas).
- **CAB:** *Chlorophyll content in the leaf* (Variáveis Biofísicas).

- **CW**: *Canopy water content* (Variáveis Biofísicas).

Nos ficheiros originais, os dados temporais, 10 ou 11 datas distintas por cidade, estavam apresentados em coluna. Assim, foi necessário adaptá-los, para termos cada uma das datas como uma instância diferente, ainda que relativa a um ponto já existente.

### 3.2.5 Dimensão dos dados

Os dados iniciais foram disponibilizados em 6 ficheiros, um por cada cidade, exceto os de Almada e Lisboa que estavam no mesmo ficheiro. A Tabela 3.1 mostra a dimensão original dos dados e a dimensão após o tratamento das datas descrito no ponto 2. da subsecção anterior.

Tabela 3.1: Dimensão dos dados.

Instâncias por categoria antes da agregação		
Broad_deciduous	331 748	66.38%
Grassland	172 200	34.46%
Artificial	87 634	17.54%
Coniferous	37 779	7.56%
Shrubs	21 494	4.30%
Bare-Lands	19 463	3.89%
Broad_ever	19 383	3.88%
Allotment	7 031	1.41%
Lakes_and_pounds	5 500	1.10%
Rivers_and_streams	701	0.14%
Riparian_vegetation	4 919	0.98%
Flower-Bed	31	0.01%
Bambu	20	0.00%
(1)	(2)	

Instâncias por categoria após a agregação		
Broad_deciduous	331 748	66.38%
Grassland	172 200	34.46%
Artificial	87 634	17.54%
Coniferous	37 779	7.56%
Shrubs	21 494	4.30%
Bare-Lands	19 463	3.89%
Broad_ever	19 383	3.88%
Allotment	7 031	1.41%
Water_surfaces	6 191	1.24%
Riparian_vegetation	4 919	0.98%
Ornamental_Bed	51	0.01%
(1)	(2)	

Notas:

- (1) A soma das instâncias por categoria é maior que o total de instâncias (499 739) porque cada instância pode pertencer a várias categorias.
- (2) As percentagens foram calculadas sobre o total de instâncias.

## 3.3 Classificação com floresta aleatória

Nesta secção será apresentado o processo de criação do modelo de classificação que utiliza o algoritmo floresta aleatória, bem como as técnicas utilizadas para lidar com o *overfitting* e o desequilíbrio de classes.



### 3.3.1 O classificador: *RandomForestClassifier*

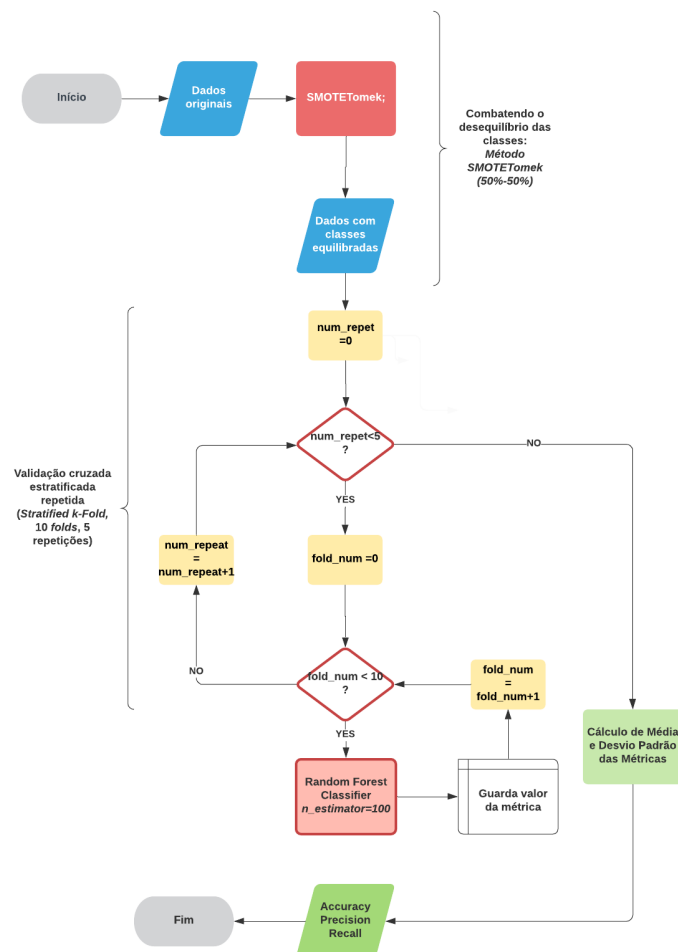


Figura 3.3: Esquema do processo associado ao classificador com floresta aleatória.

A Figura 3.3 mostra um esquema da construção do modelo. Foi utilizado um método *ensemble*, o *RandomForestClassifier*, algoritmo de classificação da biblioteca *scikit-learn* do *Python* para obtermos uma classificação binária para cada uma das 11 classes. O parâmetro *n\_estimators*, que corresponde ao nº de árvores de decisão que compõem o *ensemble*, foi definido para 100. Para entrada do modelo foi utilizada uma réplica do conjunto de dados em que os valores originais das categorias, em percentagem, foram transformados numa variável binária (os valores em que a percentagem era superior a 0 foram substituídos por 1). Para avaliação do modelo utilizámos as métricas consideradas apropriadas

para classificação binária: *accuracy*, *precision* e *recall*.

### 3.3.2 *Overfitting*: validação cruzada estratificada repetida

Para combater o *overfitting* foi efetuada uma validação cruzada estratificada repetida recorrendo ao *cross-validation score*, que usa o método *StratifiedKfold* que tem a vantagem de combater o desequilíbrio entre classes, com 10 *folds* e com 5 repetições. Desta forma, para chegar aos resultados finais das métricas escolhidas, o algoritmo vai guardando o resultado em cada uma das 50 (5 repetições x 10 *folds*) iterações. No final são apresentados a média e o desvio padrão para cada uma das métricas escolhidas.

### 3.3.3 Desequilíbrio das classes

Feita uma análise prévia dos dados, concluímos existir um desequilíbrio na quase totalidade das categorias em análise. Considerando uma das categorias de cada vez, a maioria das nossas instâncias que, recordando, correspondem a um quadrado de 10 m x 10 m representado pelas coordenadas geográficas do seu centróide, apresenta o valor 0, significando que aquela categoria não está presente naquele espaço, ou seja é uma instância negativa. A Tabela 3.2 mostra que apenas as classes *Broad Deciduous* e *Grassland* apresentam algum equilíbrio. No caso da primeira há mais instâncias positivas do que negativas. Nas restantes, o nº de instâncias positivas é escasso, havendo mesmo uma categoria que apenas surge em 0.01% dos espaços (*Ornamental bed*).

Tabela 3.2: Desequilíbrio das categorias.

Instâncias positivas por categoria.		
Broad Deciduous	331 748	66.38%
Grassland	172 200	34.46%
Artificial	87 634	17.54%
Coniferous	37 779	7.56%
Shrubs	21 494	4.30%
BareLands	19 463	3.89%
Broad Ever	19 383	3.88%
Allotment	7 031	1.41%
Water Surfaces	6 191	1.24%
Riparian Veget	4 919	0.98%
Ornamental Bed	51	0.01%
Total de Exemplos	499 739	

Para lidar com este desequilíbrio foi aplicado *SMOTE* e/ou *SMOTETomek*, neste último caso com várias proporções (80-20, 70-30, 60-40, 50-50) com resultados claramente satisfatórios. Do *GridSearch*

efetuado, o *SMOTETomek 50-50* foi o que apresentou melhores resultados para todas as classes. Na Tabela 3.3 e a título de exemplo, é visível a melhoria muito significativa na taxa de verdadeiros positivos (*Recall*), que resultou da aplicação de *SMOTETomek* à categoria *Allotment*.

Tabela 3.3: Resultado da aplicação de técnicas de correção de desequilíbrio de classes à categoria *Allotment*.

Categoria: Allotment	Método aplicado				
	Nenhum	SMOTETomek 80_20	SMOTETomek 70_30	SMOTETomek 60_40	SMOTETomek 50_50
Total Instâncias	499 739	615 613	703 756	821 152	985 406
Instâncias (-)	492 708	492 572	492 652	492 694	492 703
Instâncias (+)	7 031	123 041	211 104	328 458	492 703
% (+)	1.41%	19.99%	30.00%	40.00%	50.00%
Accuracy	97.49	96.62	97.13	97.59	98.02
Precision	70.66	92.22	94.26	95.72	96.86
Recall	17.41	96.81	98.8	99.54	99.84

## 3.4 Classificação com RNC

### 3.4.1 Pré-processamento dos dados

As RNC são tipicamente utilizadas para treinar um classificador de imagens. Para isso, são normalmente utilizados grandes conjuntos de imagens de tamanho idêntico e esse tamanho (comprimento x largura) é um dos parâmetros da entrada da rede, bem como o número de canais (que numa simples imagem a cores, RGB, são 3). Assim, teríamos, tomando como exemplo uma imagem de 10x10, uma entrada de [10, 10, 3] (largura, altura, n° de canais).

No nosso caso, a abordagem do problema através deste método enfrenta um desafio suplementar que é a transformação dos dados de forma a que possam adquirir o formato que pretendemos para a entrada da RNC. Resolvemos este desafio em várias etapas:

1. Não é possível usar cada ponto como uma instância, como fizemos para a abordagem com floresta aleatória: embora os pontos correspondam a espaços de 10m x 10m, apenas dispomos de um valor das características para cada espaço e devíamos ter  $10 \times 10 = 100$  valores (vezes o número de canais, que no nosso caso são 27). Como solução para este problema decidimos juntar vários pontos contíguos para formar imagens. Sabendo que os pontos estavam agrupados por parques dentro de cada cidade, utilizámos essa informação para agregar todos os pontos de cada parque. Para

definirmos a altura e a largura da imagem usámos, naturalmente, as coordenadas geográficas de cada ponto.

2. Como ilustra a Tabela 3.4, ficaríamos assim apenas com 266 imagens. Por outro lado, essas imagens seriam quase todas de diferentes tamanhos, algumas enormes e outras ínfimas e com um problema ainda maior relacionado com a classe que iríamos atribuir a cada imagem. O que parece mais óbvio é atribuir a cada imagem assim obtida, a classe mais representada no conjunto de pontos que a constituem, mas com imagens muito grandes perdemos detalhe e diversidade. Decidimos então dividir cada imagem (correspondendo a cada parque) em imagens de 10x10 e atribuir-lhe a classe mais vezes presente nos 100 pontos <sup>1</sup> que a constituem.

Tabela 3.4: Parques por cidade.

Cidade	Parques
Almada	16
Antuérpia	36
Lisboa	39
Paris	50
Poznan	36
Tartu	36
Zurique	53
Total	266

3. Como referimos no ponto anterior, para uma imagem de 10x10 deveríamos ter 100 pontos, no entanto isso não acontece porque os parques não são retângulos perfeitos e por isso faltam muitos pontos em zonas perto dos seus limites geográficos, como ilustrado na Figura 3.4, destacado a amarelo. Foram então adicionados pontos artificialmente, com os valores das *features* preenchidos com -1, uma vez que os valores reais tinham sido normalizados para valores entre 0 e 1, por forma a que o modelo não aprendesse com estes pontos.

---

<sup>1</sup>veremos de seguida que poderão ser menos

	1945923	1945943	1945963	1945983	1946003	1946023	1946043	1946063	1946083	1946103
2661583		0.20784	0.32941	0.38824	0.24706	0.21569				
2661603		0.29412	0.36471	0.39216	0.34118					
2661623		0.56078	0.58431	0.28235	0.29412					
2661643		0.19608	0.29020	0.33725	0.21569					0.07059
2661663	0.25882	0.47843	0.74902	0.16863	0.20000					0.04314
2661683	0.32941	0.47451	0.63529	0.23529	0.19608	0.16471				0.13333
2661703		1.00000	0.43137	0.31765	0.17255	0.17647			0.06667	0.05490
2661723		0.35686	0.32941	0.21961	0.15686	0.14902			0.03922	0.03529
2661743			0.11373	0.20784	0.06667	0.11373	0.09412	0.09412		
2661763				0.12549	0.08235	0.08235				

Figura 3.4: Exemplo de imagem 10x10 com pontos em falta.

- Adicionalmente, para aumentar o número de instâncias para aprendizagem (o número de imagens) e também para incrementar a diversidade de categorias nas imagens, aplicámos um processo semelhante ao de *stride* utilizado pela RNC e descrito no capítulo **Enquadramento e trabalho relacionado**, seleccionando imagens a começar de 5 em 5 pontos em vez de 10 em 10. Desta forma, os mesmos pontos ficam representados em várias imagens.

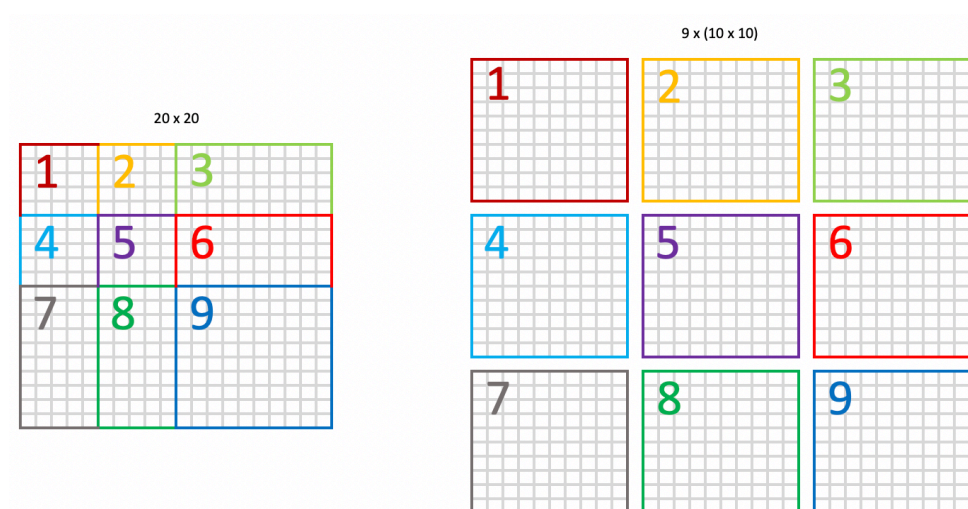


Figura 3.5: Processo que incrementa o número de instâncias para entrada da rede convolucional.

- Similarmente ao efetuado no ponto anterior utilizámos um processo semelhante ao de *padding* descrito no capítulo **Enquadramento e trabalho relacionado** para adicionar linhas e/ou colunas nos parques cuja altura ou largura não fossem múltiplas de 5. Este processo está esquematizado na Figura 3.6.

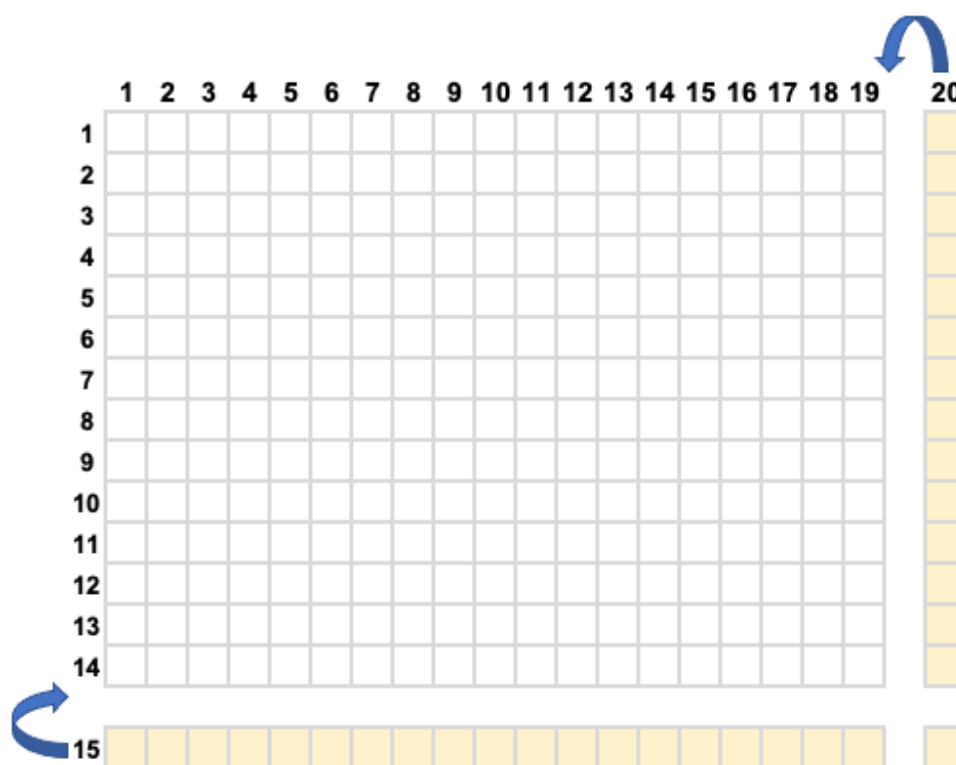


Figura 3.6: Processo para adicionar linhas e/ou colunas aos parques de forma a que a largura e altura sejam múltiplos de 5.

6. No processo de tratamento dos parques referido nos pontos anteriores, para evitar a inclusão de muitas linhas ou colunas artificiais (que ficariam com valor  $=-1$ ) nos parques muito pequenos, foram apenas considerados os parques com largura  $> 9$  e  $(\text{altura} + \text{largura}) > 16$  ou altura  $> 9$  e  $(\text{altura} + \text{largura}) > 16$ . Desta forma, garantimos que em cada parque seriam adicionadas artificialmente 2 linhas e 1 coluna ou 1 linha e 2 colunas, no máximo.
7. Foram depois excluídas as imagens em que todos os pontos estavam preenchidos com -1, resultantes dos processos descritos nos pontos anteriores. Esta situação ocorreu em parques grandes e/ou com forma muito recortada em que uma parte do retângulo não corresponde a uma área do parque.

Concluído este processo, obtivemos 32 651 instâncias, correspondendo a imagens 10x10 extraídas de todos os parques das cidades.

A atribuição da categoria a cada uma das imagens foi feita de forma distinta para a classificação multi-classe e para a classificação binária:

1. **Classificação multi-classe:** Foi atribuída a categoria com o maior valor da soma dos 100 pontos

que estiveram na sua origem. Para este processo foram utilizados os valores originais em percentagem de forma a ser considerada, não apenas a presença ou ausência daquela classe, mas também o seu peso.

O resultado deste processo foi um conjunto de dados constituído por 32 651 instâncias classificadas numa ou várias das categorias definidas, neste caso apenas 9, porque para duas delas não resultou nenhuma imagem, ou por nunca ser a categoria mais frequente ou por o tamanho da imagem ser inferior ao mínimo referido atrás. No capítulo **Apresentação e discussão dos resultados** serão apresentadas as características do conjunto de dados, em termos de distribuição das instâncias por classe, na Tabela 4.3.

2. **Classificação binária:** Foi considerada apenas a presença ou ausência da categoria em cada imagem. Desta forma, várias categorias podem ser atribuídas à mesma imagem ou, de outra forma, a mesma imagem é utilizada em mais do que um classificador, uma vez que temos um classificador para cada categoria. O resultado deste processo está resumido na Tabela 3.5. De notar que a exemplo da classificação multi-classe, também não dispomos de nenhuma imagem para a categoria *Ornamental bed*. Isto acontece, porque nos dados iniciais esta categoria apenas estava representada em 51 pontos e todos eles faziam parte de parques que tinham uma dimensão abaixo do mínimo que considerámos, como descrito no ponto 6.

Tabela 3.5: Entrada da RNC-Classificação binária-Imagens por categoria.

Categorias	Classes		
	0	1	%
Allotment	30 914	1 737	5%
Artificial	22 452	10 199	31%
BareLands	29 097	3 554	11%
Grassland	24 676	7 975	24%
WaterSurfaces	29 402	3 249	10%
RiparianVeget	29 897	2 754	8%
Shrubs	24 293	8 358	26%
BroadDeciduous	29 333	3 318	10%
BroadEver	28 810	3 841	12%
Coniferous	20 549	12 102	37%

O tratamento dos dados ficou completo com a criação de 2 *numpy arrays*:

1. Um *array* com dimensão (32 651, 10, 10, 27), onde 32 651 é o nº de instâncias (imagens), 10 a largura da imagem, 10 a altura da imagem e 27 corresponde às nossas variáveis explicativas que aqui são equivalentes a canais.

2. Outro *array* com dimensão (32 651, 1), onde para cada imagem, temos a sua classe (número inteiro correspondente à categoria). No caso da classificação binária, temos 10 *arrays* (11-1 categorias, uma vez que para a *Ornamental bed* não temos nenhuma instância).

### 3.4.2 Classificação multi-classe

O processo utilizado para a classificação multi-classe com RNC difere do processo da floresta aleatória essencialmente no pré-processamento dos dados que foram transformados em *numpy arrays* representando imagens para entrada do modelo, como é visível na Figura 3.7.

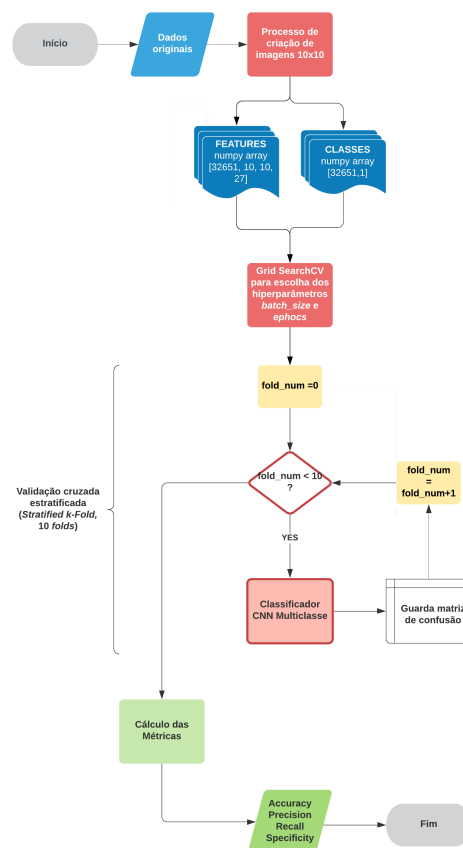


Figura 3.7: Esquema do processo associado ao classificador com RNC.

Quanto ao classificador, o *TensorFlow* e a *Keras Sequential API* permitem criar uma rede convolucional com poucas linhas de código (Figura 3.8).



```

+-----+
| # Define the model architecture |
| model = Sequential()           |
| model.add(Conv2D(32, kernel_size=(3, 3), activation='relu', input_shape=(10, 10, 27))) |
| model.add(MaxPooling2D(pool_size=(2, 2))) |
| model.add(Conv2D(64, kernel_size=(3, 3), activation='relu')) |
| model.add(MaxPooling2D(pool_size=(2, 2))) |
| model.add(Flatten())           |
| model.add(Dense(256, activation='relu')) |
| model.add(Dense(128, activation='relu')) |
| model.add(Dense(11, activation='softmax')) |
+-----+

```

Figura 3.8: RNC multi-classe: Definição da arquitetura do modelo.

Assim, criamos um modelo sequencial com duas camadas convolucionais, duas camadas *pooling* e três camadas densamente conectadas. Na primeira camada convolucional é definida a forma da entrada (imagens 10 x 10, com 27 canais). A última camada densamente conectada define o tamanho da saída, neste caso 11, que corresponde ao número de categorias que podem ser atribuídas à imagem.

Para as camadas convolucionais ensaiamos filtros (3x3), (5x5) e (7x7), como veremos no capítulo **Apresentação e discussão dos resultados**.

Para mitigar o risco de *Overfitting*, a exemplo do que fizéramos na abordagem com floresta aleatória, utilizamos validação cruzada estratificada (*stratifiedKFold*) com 10 *folds*.

Para definição dos hiperparâmetros *batch\_size* e *epochs* utilizamos o método *GridSearchCV*, como veremos em detalhe na última secção deste capítulo.

### 3.4.3 Classificação binária com um classificador RNC por categoria

Os passos seguidos nesta abordagem não diferem muito dos da anterior, a não ser que em vez de um único modelo, temos um modelo sequencial para cada categoria, que apenas difere do da multi-classe na camada de saída, que tem o valor 1, e utiliza a função de ativação *sigmoid* (Figura 3.9).

```

+-----+
| # Define the model architecture |
| model = Sequential()           |
| model.add(Conv2D(16, kernel_size=(3, 3), activation='relu', input_shape=(10,10,27))) |
| model.add(MaxPooling2D(pool_size=(2, 2))) |
| model.add(Conv2D(32, kernel_size=(3, 3), activation='relu')) |
| model.add(MaxPooling2D(pool_size=(2, 2))) |
| model.add(Flatten())           |
| model.add(Dense(512, activation='relu')) |
| model.add(Dense(1, activation='sigmoid')) |
+-----+

```

Figura 3.9: RNC binária: Definição da arquitetura do modelo para cada categoria.

Foi efetuada a seleção dos parâmetros *batch\_size* e *epochs*, através do *GridSearchCV*, para cada um dos classificadores.

### 3.4.4 Escolha de parâmetros: *GridSearchCV*

Os parâmetros *batch\_size* e *epoch* são importantes quer para o desempenho do classificador quer em termos de custo computacional. Mas não existe uma regra científica para a sua otimização que varia

muito em função do tipo e natureza dos dados. Por isso, a aplicação de um método de *GridSearchCV* é indubitavelmente uma abordagem adequada. Ainda assim, como as possibilidades de escolha de combinações destes dois parâmetros são praticamente infinitas, é razoável verificar apenas os valores dos parâmetros que empiricamente são considerados adequados à generalidade das situações.

Assim, utilizámos o método *GridSearchCV* da *sklearn*, quer para a classificação RNC multi-classe quer para a classificação RNC binária para testar as combinações entre *batch\_size* = [10, 20, 30, 40, 50] e *epochs* = [10, 25, 50, 100]. No caso do classificador multi-classe os valores resultantes foram, respetivamente 50 e 100. Para os classificadores binários resultaram os valores constantes da Tabela 3.6. Estes foram então, os valores utilizados para o processo de aprendizagem em cada um dos modelos.

Categorias	batch_size	epochs
Allotment	10	50
Artificial	20	100
Bare lands	30	100
Broad deciduous	30	50
Broad ever	30	50
Coniferous	30	50
Grassland	40	100
Riparian vegetation	40	100
Shrubs	40	50
Water surfaces	50	100

Tabela 3.6: Resultados do *Grid Search* para a RNC binária.

# Capítulo 4

## Apresentação e discussão dos resultados

Nas três primeiras secções deste capítulo são apresentados os resultados de cada uma das três abordagens de aprendizagem automática utilizadas. Na última secção são analisados e discutidos os resultados obtidos nas três abordagens.

### 4.1 Classificação binária com floresta aleatória

Nesta secção apresentam-se os resultados obtidos pela aplicação da metodologia descrita na Secção 3.3 do Capítulo **Metodologia**. Recordando essa descrição, foram utilizadas 11 réplicas do conjunto inicial de dados, cada um constituído por 499 739 instâncias, correspondentes a espaços de 10 metros por 10 metros, representados pelas coordenadas geográficas do seu centróide. Cada instância inclui 27 variáveis explicativas e a classificação binária relativa à categoria a tratar, com as 2 classes distribuídas da forma que pode ser visualizada na parte direita da Tabela 4.1.

Tabela 4.1: Resultados da classificação binária com floresta aleatória.

Categorias	Métricas			Inst.c/ classe=0 (Ausente)	Inst.c/ classe=1 (Presente)	
	Accuracy	Precision	Recall			
Ornamental_Bed	99.99 (+-0.01)	99.99 (+- 0.01)	100.00 (+- 0.01)	499 688	51	0.01%
Allotment	98.02 (+-4.13)	96.86 (+- 6.41)	99.84 (+- 0.02)	492 708	7031	1.41%
Riparian_Veget	97.75 (+-2.60)	96.04 (+- 4.55)	99.86 (+- 0.09)	494 820	4919	0.98%
WaterSurfaces	97.23 (+-2.01)	95.13 (+- 3.48)	99.75 (+- 0.05)	493 548	6191	1.24%
Bare_Lands	94.80 (+-7.99)	92.65 (+-10.22)	99.20 (+- 0.07)	480 276	19 463	3.89%
Broad_Ever	93.52 (+-7.91)	90.59 (+-10.40)	99.13 (+- 0.08)	480 356	19 383	3.88%
Shrubs	90.80 (+-6.98)	86.24 (+- 9.48)	98.86 (+- 0.15)	478 245	21 494	4.30%
Coniferous	88.17 (+-8.99)	83.77 (+-11.05)	97.57 (+- 0.30)	461 960	37 779	7.56%
Artificial	79.02 (+-8.26)	73.42 (+- 9.00)	93.81 (+- 3.05)	412 105	87 634	17.54%
Broad_Deciduous	78.79 (+-6.81)	88.88 (+- 7.46)	66.60 (+-13.74)	167 991	331 748	66.38%
Grassland	70.28 (+-7.23)	65.92 (+- 6.99)	86.42 (+- 6.87)	327 539	172 200	34.46%
Total de Instâncias				499 739		

Os resultados obtidos não indiciam que este modelo seja generalizável porque, em particular para as categorias mais bem representadas, as métricas utilizadas para a sua avaliação apresentaram valores modestos. Tomando como exemplo a categoria *Grassland*, a *accuracy* não foi além de 70%, como evidencia a Tabela 4.1. Em sentido contrário, mas com o mesmo racional, os resultados com *accuracy* mais alta ocorrem nas categorias com menor frequência, e.g. *Ornamental bed* em que todas as instâncias foram bem classificadas, mas tendo em conta que eram apenas 51, não sabemos o que aconteceria em presença de um conjunto de dados maior e com maior diversidade de características.

Os desempenhos aqui apresentados são já resultado da aplicação das técnicas de correção do desequilíbrio de classes. A Tabela 4.2 mostra como a aplicação do método *SMOTETomek* permitiu melhorar o desempenho do classificador, para a categoria *Coniferous* aqui apresentada como exemplo. Esta melhoria foi particularmente significativa no acréscimo de *verdadeiros positivos*, visível nas métricas *precision* e *recall* que passaram, respetivamente de 42.0% para 83.8% e de 16.5% para 97.6%.

Tabela 4.2: Classificação binária - *SmoteTomek* para a categoria *Coniferous*.

Categoria: Coniferous	Método aplicado				
	Nenhum	SMOTETomek 80_20	SMOTETomek 70_30	SMOTETomek 60_40	SMOTETomek 50_50
Total Instâncias	499 739	574 240	658 341	768 703	923 042
Instâncias (-)	461 960	460 355	461 159	461 345	461 521
Instâncias (+)	37 779	113 885	197 182	307 358	461 521
% (+)	7.56%	19.83%	29.95%	39.98%	50.00%
Accuracy	86.33	84.44	85.10	86.50	88.17
Precision	42.07	67.97	74.74	79.70	83.77
Recall	16.53	76.48	89.33	94.88	97.57

Embora se apresentem aqui apenas os resultados da técnica *SMOTETomek* para a categoria *Coniferous* a técnica foi aplicada a todas as categorias, sendo que, em todas elas o rácio 50/50 foi o que obteve melhor resultado.

## 4.2 Classificação multi-classe com RNC

O conjunto de dados utilizados nesta abordagem corresponde a 32 651 instâncias obtidas no processo descrito na Secção 3.4. Apenas 9 categorias estão presentes na entrada, uma vez que duas das categorias, *Ornamental bed* e *Water surfaces* não ficaram representadas após o processo de criação das imagens. A primeira, porque o tamanho dos parques onde está referenciada é inferior ao mínimo considerado, como descrito na metodologia, [largura > 9 e (altura+largura) >16] ou [altura >9 e (altura+largura) >16]. A segunda, porque em nenhum dos parques onde está referenciada é a categoria mais frequente. Na Tabela 4.3 podem ser visualizadas as características deste conjunto de dados.

Tabela 4.3: Classificação multi-classe-Imagens por categoria.

Categorias	Freq.	Freq. %
Allotment	590	1.8%
Artificial	807	2.5%
Ornamental_Bed	0	0.0%
Bare_Lands	880	2.7%
Grassland	7 869	24.1%
Water_Surfaces	0	0.0%
Riparian_Veget	55	0.2%
Shrubs	132	0.4%
Broad_Deciduous	19 959	61.1%
Broad_Ever	1 110	3.4%
Coniferous	1 249	3.8%
<b>Total</b>	32651	

Utilizando um único classificador RNC foram obtidos excelentes resultados como é visível na Tabela 4.4. Foram testados três diferentes tamanhos de filtros, não se verificando diferença significativa entre os resultados obtidos. Em todos os casos obteve-se *Accuracy* superior a 97%.

Tabela 4.4: Resultados da classificação multi-classe com RNC.

Tamanho dos filtros	Accuracy	Loss
3x3	97.81 (+- 0.28)	0.07
5x5	97.77 (+- 0.47)	0.07
7x7	97.26 (+- 1.05)	0.08

Para melhor avaliar o desempenho do classificador vamos aqui utilizar matrizes de confusão que permitem visualizar em que categorias ou classes o modelo foi mais ou menos eficiente. A Figura 4.1 apresenta, para cada categoria, os valores da previsão (em coluna) contra os valores reais (em linha). Podemos, por exemplo, constatar que em alguns casos, houve alguma dificuldade em distinguir relvados, *Grassland* de árvores de folha larga caduca, *Broad\_deciduous*, uma vez que 264 espaços compostos maioritariamente por relva foram classificados como árvores de folha caduca. Pelo contrário, os espaços pré-classificados como terra nua, *Bare Lands*, foram quase todos bem classificados pelo modelo, apenas 3 em 880, (0.34%), não o foram.

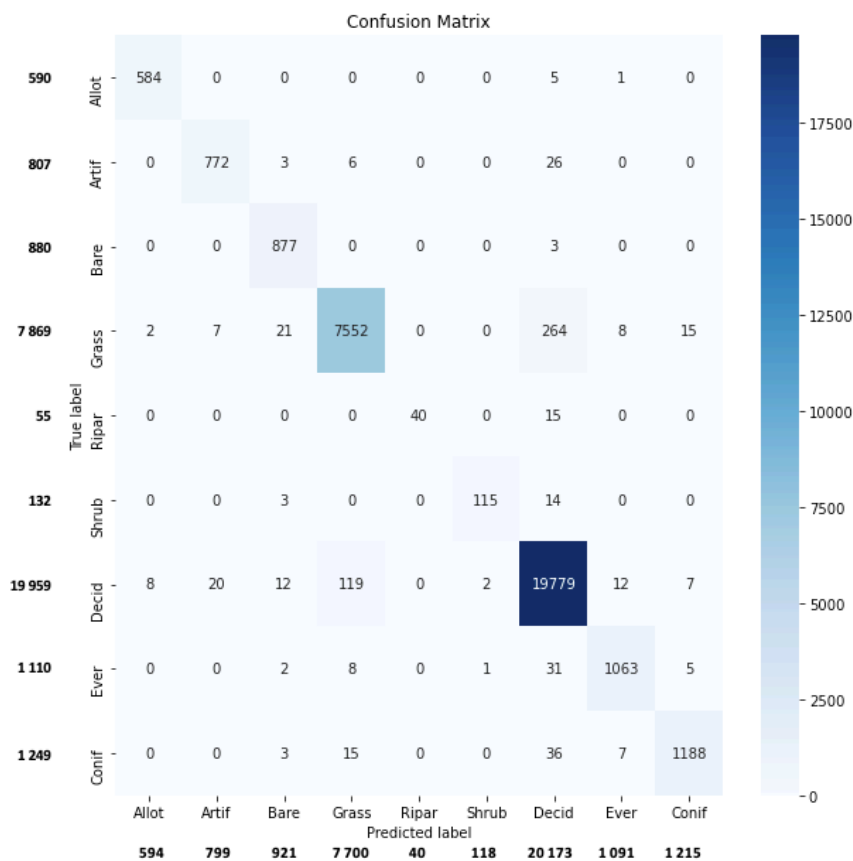


Figura 4.1: RNC multi-classe: Matriz de confusão com os valores reais vs valores previstos, para cada categoria.

A matriz de confusão da Figura 4.2 apresenta os valores da matriz inicial divididos pelo número de instâncias total da categoria a que realmente pertencem, permitindo assim entender que percentagem das instâncias pertencentes a uma categoria foi classificada nessa ou noutra categoria. De destacar que a categoria *Riparian vegetation* é a que apresenta o pior desempenho, com 27% das instâncias a serem incorretamente classificadas como *Broad deciduous*, embora tendo em conta que aquela categoria apenas

tem 55 instâncias positivas. Ao contrário, na categoria *Bare-lands* o valor 1.00 indica que quase todas as instâncias desta categoria foram corretamente classificadas ( $\frac{877}{880}$ ). Comparando com a matriz anterior é de realçar que algumas categorias como *Broad ever* e *Coniferous* que estavam destacadas por apresentarem valores elevados, em termos percentuais apresentam valores relativamente mais baixos, 96% e 95%, respetivamente.

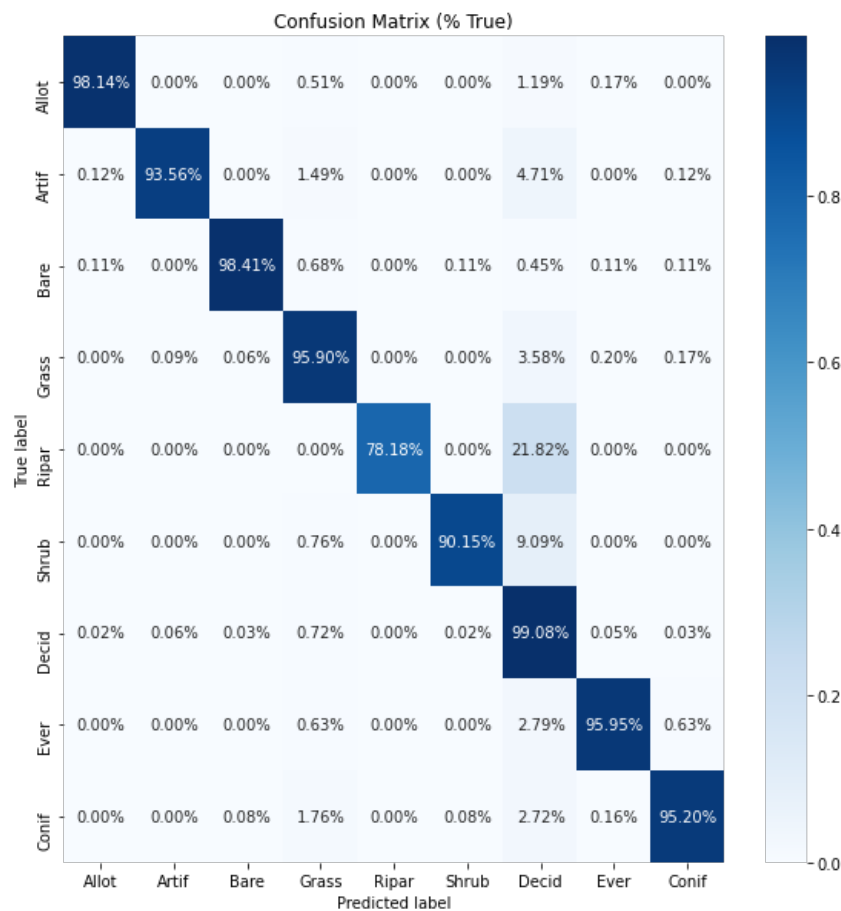


Figura 4.2: RNC multi-classe: Matriz de confusão normalizada com os valores previstos divididos pelo valor real total de cada categoria.

Outra forma de apresentar os resultados através de uma matriz de confusão normalizada é dividir cada elemento da matriz original pelo total de instâncias previstas para a categoria a que realmente pertencem (Figura 4.3). Por exemplo, podemos ver que todas as instâncias classificadas como *Riparian vegetation* pertencem realmente a essa classe. Ao contrário, na categoria *Bare-lands* apenas 95% das instâncias classificadas como tal, o são realmente.

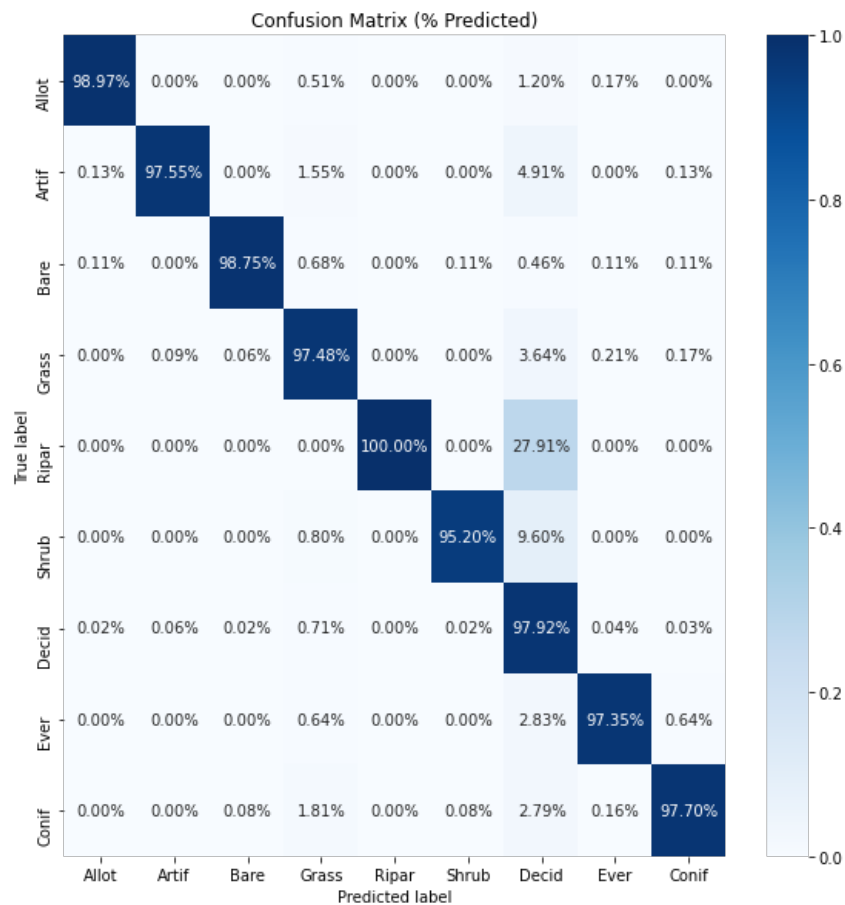


Figura 4.3: RNC multi-classe: Matriz de confusão normalizada com os valores previstos divididos pelo valor previsto total de cada categoria.

### 4.3 Classificação binária com RNC

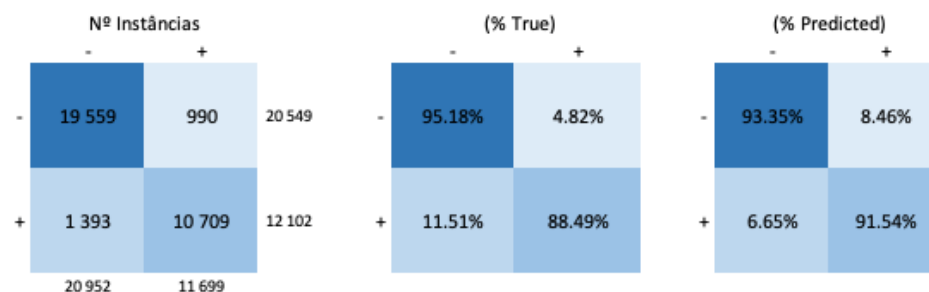
Nesta última abordagem foram utilizados 10 classificadores RNC binários, um para cada categoria, excluindo a classe *Ornamental bed* para a qual não existem observações suficientes para criar uma imagem, pois todos os espaços em que esta categoria estava identificada eram inferiores ao mínimo considerado para criar imagens. Recorde-se que para este modelo foram criadas 9 réplicas das matrizes correspondentes às 32 651 imagens criadas, à semelhança do modelo RNC multi-classe, mas neste caso, cada uma delas está classificada com 0 e 1, correspondendo à ausência ou presença da categoria na imagem. A parte direita da tabela 4.5 mostra a distribuição das classes por cada categoria.



Tabela 4.5: Resultados da classificação binária com rede neuronal convolucional.

Categorias	Métricas				Inst.c/ classe=0 (Ausente)	Inst.c/ classe=1 (Presente)	
	Accuracy	Precision	Recall	Specificity			
Allotment	99.70 (+ 0.17)	98.87 (+ 1.34)	95.45 (+ 2.35)	99.94 (+ 0.07)	30 914	1 737	5%
Bare_Lands	99.46 (+ 0.24)	97.13 (+ 1.24)	97.95 (+ 1.37)	99.65 (+ 0.15)	29 097	3 554	11%
WaterSurfaces	99.23 (+ 0.36)	97.99 (+ 1.59)	94.18 (+ 2.78)	99.79 (+ 0.17)	29 402	3 249	10%
Broad_Deciduous	99.15 (+ 0.13)	99.27 (+ 0.19)	99.79 (+ 0.11)	93.52 (+ 1.73)	3 318	29 333	90%
Broad_Ever	99.04 (+ 0.62)	96.49 (+ 4.15)	95.55 (+ 2.08)	99.51 (+ 0.63)	28 810	3 841	12%
Riparian_Veget	98.90 (+ 0.87)	94.21 (+ 7.89)	93.86 (+ 4.51)	99.37 (+ 1.04)	29 897	2 754	8%
Shrubs	97.16 (+ 0.87)	94.72 (+ 2.67)	94.26 (+ 2.41)	98.16 (+ 1.01)	24 293	8 358	26%
Grassland	96.76 (+ 0.44)	98.01 (+ 0.74)	97.71 (+ 0.50)	93.83 (+ 2.37)	7 975	24 676	76%
Artificial	95.56 (+ 0.70)	97.04 (+ 1.04)	96.50 (+ 1.08)	93.48 (+ 2.43)	10 199	22 452	69%
Coniferous	92.70 (+ 3.39)	91.74 (+ 5.28)	88.49 (+ 5.76)	95.18 (+ 3.34)	20 549	12 102	37%
Total de Instâncias em cada classe					32 651		

A Tabela 4.5 mostra os bons resultados obtidos por cada um dos classificadores, com praticamente todas as métricas acima de 90%, embora seja de destacar a categoria *Coniferous* que apresenta um desempenho mais fraco, com uma *Accuracy* de 92.70% sendo mesmo a única categoria que apresenta uma taxa de verdadeiros positivos, (*recall*), abaixo de 90% (85.93%) o que significa que há muitos *Falsos Negativos*, ou seja, muitas instâncias pertencentes a esta categoria foram classificadas como não o sendo. Nas matrizes de confusão da Figura 4.4 podemos ver que 1 393 instâncias, 11.51% do total de instâncias positivas, deviam ter sido atribuídas a esta categoria e erradamente não o foram.

Figura 4.4: RNC Binária: Matrizes de confusão da categoria *Coniferous*.

Podemos também observar que o número de *Falsos Positivos* é elevado, 990 instâncias, correspondendo a 4.94% do total de instâncias negativas. Na matriz mais à direita pode verificar-se que 8.46% do total de instâncias classificadas como positivas não o são, realmente. Estas evidências estão também presente no valor relativamente baixo da métrica *Precision*, 91.74%.

Estes resultados poderão eventualmente ser explicados pelas características dos dados:

- As refletâncias das Coníferas são próximas das de outras categorias, e.g. Arbustos (*Shrubs*) ou Relvados (*Grassland*). Se observarmos as métricas obtidas para essas categorias também verificamos

desempenhos relativamente mais fracos.

- Por outro lado, a classificação no terreno foi efetuada atribuindo uma percentagem a cada um dos tipos de vegetação existente num espaço de 10 metros por 10 metros. Se frequentemente ocorrer a existência simultânea de um conjunto de categorias, pode acontecer alguma confusão entre elas.

#### 4.4 Análise e discussão dos resultados

A abordagem efetuada com floresta aleatória não demonstrou ser a mais adequada para o objetivo de classificar a cobertura do solo em EVU. Como vantagens pode referir-se o facto de não requerer um tratamento muito elaborado dos dados, permitir classificar todas as categorias propostas e além disso, o modelo tem o potencial de ser interpretável. Como desvantagem, é de referir apenas a necessidade de usar um classificador por categoria, mas os resultados obtidos foram modestos e não permitiram uma boa classificação. A baixa frequência de algumas categorias, nos dados disponíveis permitiu, recorrendo a técnicas de correção de desequilíbrio de classes, obter resultados aparentemente melhores para os classificadores dessas categorias. No entanto, esses resultados podem piorar ao adicionarmos mais dados, eventualmente com características mais diversas.

A classificação binária com RNC, apresenta como desvantagens, igualmente o facto de requerer um classificador por categoria, maior complexidade no pré-processamento dos dados e não permitir classificar todas as categorias, para os dados disponíveis. Mas os resultados obtidos foram claramente melhores, com apenas uma categoria a apresentar resultados mais fracos, mas ainda assim com uma *Accuracy* superior a 90%.

Finalmente, a classificação multi-classe com RNC posicionou-se como a abordagem mais adequada ao problema em estudo. Permitiu resultados globais muito bons, utiliza apenas um classificador que com uma única saída, permite classificar as instâncias numa das categorias presentes na entrada. A análise das matrizes de confusão sugere que as categorias menos representadas não foram tão bem classificadas como as restantes, mas esse é um problema que pode ser potencialmente resolvido com a obtenção de mais dados classificados nessas categorias.

## Capítulo 5

### Conclusões e trabalho futuro

Nesta dissertação foram propostas abordagens para análise de uso e cobertura de solo para os EVU, baseada em aprendizagem automática, a partir de imagens de satélite, utilizando um conjunto variado de características obtidas dessas imagens, as bandas de refletância do espectro visível e do espectro invisível de radiação e ainda de um conjunto de índices específicos, de solo, vegetação, água, densidade urbana e outras variáveis biofísicas, dados obtidos através do programa *copernicus*, para um conjunto de parques de 7 cidades europeias, para as quais havia também uma classificação obtida através do programa Bioveins [3]. Foi proposta uma classificação com um conjunto inicial de 13 classes, que foi reduzido no decorrer do estudo devido à baixa frequência de ocorrência de algumas delas nos dados disponíveis. Foram testados 2 algoritmos distintos de aprendizagem automática, um mais convencional, floresta aleatória e outro mais específico, RNC, popularmente utilizado para a classificação de objetos em imagens.

A rede neuronal convolucional demonstrou claramente ser mais adequada ao problema em estudo pois além de ter apresentado os melhores resultados em termos comparativos obteve um desempenho próximo dos 98%, considerando a métrica de avaliação *accuracy*.

Em termos de trabalho futuro, pensamos ser pertinente a introdução de algumas melhorias, particularmente com a obtenção de mais dados pré-classificados que alarguem o conjunto de categorias para além daquelas que acabámos por utilizar, pois algumas ficaram excluídas devido à sua baixa ocorrência.

A recolha de mais dados também poderá permitir outro tipo de abordagem, como por exemplo, utilizar dados de uma cidade para aprendizagem e testar o modelo em outras cidades.

Outra abordagem que poderá ser feita como desenvolvimento deste estudo é a criação de um modelo de classificação *multi-label*, que permita obter mais detalhe ao atribuir a um mesmo espaço a ocorrência de várias categorias, o que seria ainda mais adequado ao problema em estudo.



# Referências

- [1] A. M. Abdi. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience and Remote Sensing*, 00(00):1–20, 2019. ISSN 15481603. doi: 10.1080/15481603.2019.1650447. URL <https://doi.org/10.1080/15481603.2019.1650447>. 15
- [2] AWS. Algoritmo XGBoost, 2019. URL [https://docs.aws.amazon.com/pt\\_br/sagemaker/latest/dg/xgboost.html](https://docs.aws.amazon.com/pt_br/sagemaker/latest/dg/xgboost.html). 10
- [3] Bioveins. , 2017. URL <http://bioveins.eu>. V, VII, 19, 41
- [4] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva. Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. pages 1–11, 2015. URL <http://arxiv.org/abs/1508.00092>. 15
- [5] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique nitesh. *Journal of Artificial Intelligence Research*, 16(Sept. 28), 2002. 14
- [6] Direção Geral do Território. Atlas da União Europeia, 2016. URL <https://www.forumdascidades.pt/content/atlas-urbano-da-uniao-europeia>. Accessed: 2020-12-08. 1
- [7] Direção Geral do Território. Copernicus, 2017. URL <http://www.dgterritorio.pt/copernicus/>. 18
- [8] N. Donges. 4 REASONS WHY DEEP LEARNING AND NEURAL NETWORKS AREN ’ T ALWAYS THE RIGHT. 2019. 8
- [9] ESA-Copernicus. sentinel-2→ DATA ACCESS AND PRODUCTSSpectral signatures, 2015. URL [https://sentinels.copernicus.eu/documents/247904/1848117/Sentinel-2\\_Data\\_Products\\_and\\_Access](https://sentinels.copernicus.eu/documents/247904/1848117/Sentinel-2_Data_Products_and_Access). 18
- [10] ESA-RemoteSense. O que é a detecção remota?, 2019. URL [http://www.esa.int/SPECIALS/Eduspace\\_PT/SEM9S365P1G\\_0.html](http://www.esa.int/SPECIALS/Eduspace_PT/SEM9S365P1G_0.html). 5

- [11] ESA-SpectralSign. Spectral signatures, 2019. URL [https://www.esa.int/SPECIALS/Eduspace\\_EN/SEMPNQ3Z20F\\_0.html](https://www.esa.int/SPECIALS/Eduspace_EN/SEMPNQ3Z20F_0.html). 6
- [12] Field Cady. *The Data Science Handbook*. John Wiley & Sons, Inc., 1st edition, 2017. ISBN 9781119092933. 14
- [13] Han, Jiawei and Kamber, Micheline and Pei, Jian. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2011. ISBN 0123814790, 9780123814791. 6, 7, 10
- [14] S. Haykin. *Neural Networks - A Comprehensive Foundation* - Simon Haykin.pdf, 1999. 8
- [15] P. Helber, B. Bischke, A. Dengel, and D. Borth. Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2018-July:204–207, 2018. ISSN 1939-1404. doi: 10.1109/IGARSS.2018.8519248. 15
- [16] Infopedia. reflectância in Dicionário infopédia da Língua Portuguesa [em linha]. Porto: Porto Editora, 2003-2019. [consult. 2019-12-01 16:44:26]. Disponível na Internet:, 2019. URL <https://www.infopedia.pt/dicionarios/lingua-portuguesa/reflectância>. 6
- [17] D. Kumar. Top 4 advantages and disadvantages of Support Vector Machine or SVM Advantages :. pages 3–4, 2019. 7
- [18] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5): 778–782, May 2017. 15
- [19] Lujing Chen. Basic Ensemble Learning (Random Forest, AdaBoost, Gradient Boosting)- Step by Step Explained. *Towardsdatascience.Com*, 2019. URL <https://towardsdatascience.com/basic-ensemble-learning-random-forest-adaboost-gradient-boosting-step-by-step-explained-95d49d1e2725>. 9, 10
- [20] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez. Fully convolutional neural networks for remote sensing image classification. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, volume 2016-Novem, 2016. ISBN 9781509033324. doi: 10.1109/IGARSS.2016.7730322. 15
- [21] Maria Thomas. Neural Networks: Advantages and Applications, 2019. URL <https://www.marktechpost.com/2019/04/18/introduction-to-neural-networks-advantages-and-applications/>. 8

- [22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 12
- [23] O. A. Penatti, K. Nogueira, and J. A. Dos Santos. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 44–51, 2015. 15
- [24] SAS. Redes Neurais o que são e qual sua importância?, 2019. URL [https://www.sas.com/pt\\_br/insights/analytics/neural-networks.html](https://www.sas.com/pt_br/insights/analytics/neural-networks.html). 8
- [25] I. Tomek. Two modifications of cnn. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6 (11), 1976. 14
- [26] United Nations. 68% of the world population projected to live in urban areas by 2050, says UN, 2018. URL <https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html>. 1
- [27] Y. Yang and S. Newsam. Bag-of-visual-words and spatial extensions for land-use classification. In *ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS)*, 2010. 15
- [28] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola. *Dive into Deep Learning*. 2020. <https://d2l.ai>. 10